

# SEQUENCE LISTING

## SEQ ID NO: 1 BLM gene cluster ORFS 30 through 8

(note orf 31-40 on sequence 1-18660 are translated on the reverse strand and on a separate file)

18601	ACCCATCTCATAGGTGTACGCGCTGGAGCATTTCGGGGCACGACGGAAGGTTCTCGGTCAC	18660
18661	GAGAGCACTGTAAGCCCGAACCCGCAAGCATGACGAATTGCAAAATTGTGCAAGTCGCTA	18720
18721	CATGATGGTCCGGCTGTGCCCGCAGGTAGCCGCGGGCACAGCACCAGACGCTGCCCTCCGC	18780
18781	GCACCGCGCGGGAGGCCCGGTGAGGCGAGAGGCTGAGGTTCCGTGCCGGTTCCGCTGTAT	18840
	M P V P L Y	(orf30)
18841	CAGGCGAAGGCCGAGTTCTTCCGGATGCTGGGGCACCCGGTCCGCATCCGCGTACTGGAG	18900
	Q A K A E F F R M L G H P V R I R V L E	
18901	CTGCTGCAGGACGGGCGGATGCCGGTGCCTGATCTGCTGGCGGCGATCGAGATCGAGCCC	18960
	L L Q D G P M P V R D L L A A I E I E P	
18961	TCGGCGCTGTCCCAGCAGCTGGCGGTGTGCGCCGCTCGGGCATCGTGACCTCCACCCGC	19020
	S A L S Q Q L A V L R R S G I V T S T R	
19021	ACGGGTTCCACGGTCGTCTACGAGCTGGCCGGTGGCGACGTGGCGGAGCTGATGTCCGCC	19080
	T G S T V V Y E L A G G D V A E L M S A	
19081	GCGCGCCGCATCCTGACCGAGATGCTCAATGGGCAGCACGAGCTGCTGGAGGAGCTGAGG	19140
	A R R I L T E M L N G Q H E L L E E L R	
19141	GAAGCCGAGGTCAAGTCCCGGTGAGCTCCCTCGCCGTCCGGGTGGGAGCCCGGTGCGTT	19200
	E A E V S A R *	
	M S S L A V R V G A R V R S	(orf29)
19201	CCGTGCTGCCACCCGCGCGACCTCGCGGGCATGGGCGCAGCCCGCGACGTGATCTAC	19260
	V L P T R A D L A G M G R S P R R D L L	
19261	TGGCCGGTCTGACCGTGGCGATCGTGCCCTGCCGCTCGCCCTCGGATTTCGGCGTCTCCT	19320
	A G L T V A I V A L P L A L G F G V S S	
19321	CCGGTCTCGGCGCGGAGGCAGGGCTGGCCACCGCGGTGGTGGCGGGCGCGCTGGCCGCGG	19380
	G L G A E A G L A T A V V A G A L A A V	
19381	TATTCGGTGGGTGCAATCTCCAGGTGTCCGGGCCCCACGGGCGCCATGACCGTGGTCTCTGG	19440
	F G G S N L Q V S G P T G A M T V V L V	
19441	TGCCCATCGTCGCCCCGTACGGCCCCGGCGGTGTCTCACGGTCGGCCTGCTCGCCGGAC	19500
	P I V A R Y G P G G V L T V G L L A G L	
19501	TGATGCTGATCGCGCTCGCCCTCGCCCGCCCGGCCGTACATGCAGTACGTGCCGGCCC	19560
	M L I A L A L A R A G R Y M Q Y V P A P	
19561	CGGTGGTGGAGGGCTTACCCTCGGCATCGCTGCGTATCGGCTTGACGAGGTGCCGA	19620
	V V E G F T L G I A C V I G L Q Q V P N	
19621	ACGCCCTGGGAGTCGCCAAGCCGAGGGGACAAGGTCTCGTCGTGACCTGGCGCGCGG	19680
	A L G V A K P E G D K V L V V T W R A V	
19681	TCGAGACCTTCGCCGGGGCGCCCAACTGGACCGCTGCCGGAAGTGGCGGCGAGCGGTGCCG	19740
	E T F A G A P N W T A A G L A A A V A A	
19741	CGGTCATGCTGACCGCGCGCGGTGGCGGCGGTCGTTCCCTTCTCCCTCCTCGCGGTGA	19800
	V M L T G A R W R P V V P F S L L A V T	
19801	CCGGTGGCACCGTCTGCCCCAGCTGTGCCACCTGGACGCGGCCCGCCGATCGGGGACC	19860
	G A T V V A Q L C H L D A A R P I G D L	
19861	TGCCCCGCGGGGCTGCCCGCCCCGTGCTGGCCTTCTGGACCTCGGAGCACTGGGCTCGC	19920
	P A G L P A P S L A F L D L G A L G S L	
19921	TGCTGGCGCCTGCCGTGGCCGTGGCGGCCCTTGCCGCGTTGGAATCGCTGCTGTCGCGT	19980
	L A P A V A V A A L A A L E S L L S A S	

19981 CCGTCGCGGACGGCATGACGGTCGGGCGAGAAGCACGACCCGGACAAGGAGCTGTTCCGGGC 20040  
       V A D G M T V G Q K H D P D K E L F G Q  
 20041 AGGGTCTCGCCAACCTGGCCGCCCCGCTGTTTCGGCGGCGTCCCGGCCACCGGCGCGATAG 20100  
       G L A N L A A P L F G G V P A T G A I A  
 20101 CCCGCACCGCCGTCAACGTCCGTACCGGTGCGAGCTCGCGACTGGCGGCCCTCACGCACG 20160  
       R T A V N V R T G A S S R L A A L T H A  
 20161 CCGCGATCCTCGCCGTATCGTCTTCGCCGCGCCCGCCACTGGTCTCCCGCATCCCCCTGG 20220  
       A I L A V I V F A A A P L V S R I P L A  
 20221 CCGCGCTCGCCGGCGTGCTGATCGCGACCGCGATCCGCATGGTCTGAAGTGGGCAGCCTGC 20280  
       A L A G V L I A T A I R M V E V G S L R  
 20281 GGGCGATGGCCCGCGCCACGCGCTCCGACGGCCTGGTACTGATCCTCACGGCGGTGCGCA 20340  
       A M A R A T R S D G L V L I L T A V A T  
 20341 CCGTGGCCCTGGACCTCGTCTACGCCGTATCATCGGCCCTGCTGGTCTGCGCGGCGCACTCG 20400  
       V A L D L V Y A V I I G L L V A G A L A  
 20401 CCCTGCGGGCCGTGGCCAAGCAGGTCCGCTGGACCAGGTCTCCTTGAAGGAGGACCTGA 20460  
       L R A V A K Q V R L D Q V S L K E D L T  
 20461 CCGGCGACCACAGCGCCGAGGAACACGCGTGTCTCGCCGAGCACATCGTGGCGTACCGCA 20520  
       G D H S A E E H A L L A E H I V A Y R I  
 20521 TCGACGGTCCGCTGTTCTTCGCCGCGGCCACCGCTTCTCTGTTGAAGTCTCGGACGTCTG 20580  
       D G P L F F A A A H R F L L E L S D V A  
 20581 CGGACGTGCGCGTGGTGATCCTGCGCATGTCCCGCGTGACCACCATGGACGCCACCGGCG 20640  
       D V R V V I L R M S R V T T M D A T G A  
 20641 CCCTCGTCTGAAGGACGCGGTACCAAGTGAACCGGCGCGGCATCACCGTCTTGGCCT 20700  
       L V L K D A V T K L N R R G I T V L A S  
 20701 CCGGGGTACGCCCCGGCCAGCGCCGGGTCTCGACTCCGTGCGCGCCCTCGGTCTGCTCC 20760  
       G V R P G Q R R V L D S V G A L G L L R  
 20761 GGGCCGCCACCGGCGACGACTACACCGGCACTCCCGAAGCCATCGCCGCGCGCCCGAAGCC 20820  
       A A T G D D Y T G T P E A I A A A R S H  
 20821 ACCTGCACGGCGCGGTGTCTGCGCCCCCGCTGCGCGGGCCCGCCTCCTCCGGTACCCC 20880  
       L H G A G V L A P A C P G P P P P V P P  
 20881 CACCGTGCCTCCGAGTGCCCGACGATGAGGAGCCGACCGAGGTCTCCTCCGTACCCG 20940  
       P C A P S A R R \*  
 20941 GACACCCACGGTTGCGCCGCCCCATGCCGCGGTCCCTCCTGACGGCCCGTCCGCGGCTT 21000  
 21001 GAGGCGGCGGTGGACGGCCTGCCGCGCGCGGCTCGGGCTGATCGGCGTGATCACCGCCC 21060  
 21061 ATGCGCGGGTGGGCGCCCGCGGCATCGTGGGCGGGACCGTGTTCGGGCCACCGCGGCGG 21120  
 21121 CCGGCCTCGCGTGGGCGTGGCCTGCCGCGGTGCTGGTAGCGGCGGGGTCCGGCGGCCG 21180  
 21181 GGCCTGTGCTTCTTCCCGCCCGTCCGGCGGTGGCGCCGCGCGCGGTGACAGGGAAAT 21240  
 21241 ATGACCGGAAGTGGGATGCTCGCGTCCACTCGGGTGTGTTTAAAGTGCCACGGGGGCTTCC 21300  
 21301 GACGGCGCGTCCGCGCGCGCGGCTTCGCCCGATGATGGTCTGCGGCGCTGTGAGCCGGG 21360  
 21361 GAGCCTATGGCACAGGACCTGAACGACTGGATCGAGGACGAGGTCTCCCTTACGAGGAG 21420  
       M A Q D L N D W I E D E V V P Y E E (orf28)  
 21421 AAGCCTCTCGAATGGATCTCCAGTACCCTTCTTCCGCGACCCGGCGCGAGCCGCCTAT 21480  
       K P L E W I S Q Y H F F R D P A R A A Y  
 21481 GTCGATCACACCTACTTCTTCTACCGGCGGATGGCGCGATCGTCTACCAGAAAGTAGTG 21540  
       V D H T Y F F S P A D G A I V Y Q K V V  
 21541 GATCCCCAGGAGTCGATCATCGACATCAAGGGGAAGCCGTACTCGCTGGCCGCGCGCTC 21600  
       D P Q E S I I D I K G K P Y S L A A A L

21601	CGTGACGAATCGTTCCGGTCACCGGTGCCTGGTGATCGGCATCTTCATGACCTTCTTCGAC	21660
	R D E S F G H R C L V I G I F M T F F D	
21661	GTGCACATCAACCGGATGCCTTACGGCGGCCGTCTCTCCTTCGCGCTCAAGGAGCCCATC	21720
	V H I N R M P Y G G R L S F A L K E P I	
21721	GGGACGTTCAACCTCCCCATGCTGGCCATGGAGCAGGACCTGCTCGAACGGCTCCGGGTC	21780
	G T F N L P M L A M E Q D L L E R L R V	
21781	AATCCGGCTCACGCGAGGTATCTGCACCTGAACGAGCGGATGGTCAACCGGGTCGACGCG	21840
	N P A H A R Y L H L N E R M V N R V D A	
21841	CCGCGGCTCCGGGGCCCGTACTGGATGCTCCAGATCGCCGACTACGACGTCGACTCCATC	21900
	P R L R G P Y W M L Q I A D Y D V D S I	
21901	ACCCCGTTCTGCAGACGGCAGGGAATGTTCCGCTCCCAGGGCGCCGCTTCTCCCAGATC	21960
	T P F C R R Q G M F R S Q G R R F S Q I	
21961	CGCTACGGATCGCAGGTGCACCTGGTGATCCCGATGGCGGCCGACCGGAGTACGTCCCC	22020
	R Y G S Q V D L V I P M A A D R E Y V P	
22021	GTGGAGGCCGTGCGCCGGCACGTGAAGGCGGGGCTCGACCCGCTCGTCAAGATCCGGTGG	22080
	V E A V G R H V K A G L D P L V K I R W	
22081	CGTTGAAGAGCGCGTACGAAGCGATGGCGAACTGGAGGGACACAGCGTGGGTTTCCGTCG	22140
	R * M G F R R (orf27)	
22141	AGCGCAGAGGGCCGTTGGGCCGGGAGCGGGCCGGCGGGAGAGCGCCCGTTTCAGGCCGA	22200
	A Q R A G G P G A G R R E S A R F R P D	
22201	CGGGCCGTGCGCGCCGCGGGACCGTCCGTTACCCCTGTCCGCGGGCAGTTGTTTCGAGTG	22260
	G P S A P R D R P L P L S A G Q L F E W	
22261	GGTGTGTTGACAAGCTCGTCGACGGAGATCTGAGCCACCAGCCGACGATTGTGCGGCTCCG	22320
	V F D K L V D G D L S H Q P T I V R L R	
22321	CGGCCCCGCTGAACACCGCCGCCCTGCGGATGGCCTACGCCCGGCTGGTGCGGCCACGA	22380
	G P L N T A A L R M A Y A R L V R R H E	
22381	GTGCCTGCGCACCCGCTTCCCCGTGATCGACGGGGAGCCCGTGCAGGTGATCGAGGGCAT	22440
	C L R T R F P V I D G E P V Q V I E G I	
22441	CGGGAAAGCAGCGGGGGCCCGCTGCCGCTCATCGATCTGCGCCACCTCCCGGAGGCGCT	22500
	G K A A G G P L P L I D L R H L P E A L	
22501	TCGCGCGCGGAGATCGCGAGGATCCGCGAGGAGACGCTGTCCACGCCGGTCCCCCTCGA	22560
	R A R E I A R I R E E T L S T P V P F D	
22561	CAAGCGGCCCGCCCGTCCGCGTGGCGCTGATCCGGGCGGCCCGAGGAGCACCTCTTCCT	22620
	K R P P V R V A L I R A A P E E H L F L	
22621	CGTCGGCATCCCGCACATCACCGCGGACCTGTGGTCCGCGACCTGCTCAACGACGAGCT	22680
	V G I P H I T A D L W S A T L L N D E L	
22681	CATGGCGCACTACAGGGCGGGGCGAGGGGACTCCCTCCCGGGCCCCACCCCGTCGC	22740
	M A H Y R A G A E G T P S R A P T P V A	
22741	GCAGTACGCCGACTTCGCGCAGTGGCAGCGCGCTGGTGAACCGGGACCGCACCGAGCG	22800
	Q Y A D F A Q W Q R A W W N R D R T E R	
22801	GGAGGCCGGACGGTGGCGGGCGGGCTGGACGGGCTGTCCGCCGTGGAAGTGGCCCTGGA	22860
	E A G R W R A R L D G L S A V E L P L D	
22861	CCGGCCCCGCCCCGCGGGCCCGCGCGGGACTGCTTCCTGATCGGGGACACCTTCGACGC	22920
	R P R P A G R R R D C F L I G D T F D A	
22921	CGAAGTGAAGCGACCGGCTGCGGCCTTGGCACGCACCGCCGACGTCACGCTGTACGTGGT	22980
	E L S D R L R A L A R T A D V T L Y V V	
22981	GCTGTGGCGGCGTTCCACTGGCTGGTGGGGCGGATGTGGGCGCCGGCCGGCTGGTGAC	23040
	L L A A F H W L V G R M S G A G R L V T	
23041	CACCTCGCTCGTGGCCGCCCGGCACGGCAGCGCGGTACAGGGGATGACCGGCCCGTTCTC	23100

T S L V A A R H G S A V Q G M T G P F S

23101 GGACTACCTGGCCCTGGTCGGGGACCTGTGCGGGCGATCCGGACTTCTGGAGTCCCTGCGG 23160  
D Y L A L V G D L S G D P D F L E S L R

23161 CCGCGTACGCGACGAGTGCCTGACCGCCACGACCACGCGGCTTCCGTTCTCACAGGT 23220  
R V R D E C L T A H D H Q R L P F S Q V

23221 CCTCGAAGTCATGGACCCCGGACGCGAGTTGCACCCCCATCCGCTGGAGCAGCTCGGGTT 23280  
L E V M D P G R E L H P H P L E Q L G F

23281 CAACCTCCACAACATCCCTCCCGCGGTTCATGGAATTCTCCGCGACGTCGTCGTCTCGGC 23340  
N L H N I P P A V M D F S G D V V V S A

23341 GGTGAACCCGAGGGGGACGACGGGGAGAGCGGCGACGGGAGTACGTGCCCTGGACCGC 23400  
V N P E G D D G E S G D G E Y V P W T A

23401 CGACCTGACCTTCGACGTCTACGACTACGGCACCGGCCATATGCCGTTTCGACGTGATACT 23460  
D L T F D V Y D Y G T G H M P F D V I L

23461 CGACCGGCGGCTGGCCGATCCGCGGACGCGCCGGGAGTGGGCGGGCACTACCGGTCGGT 23520  
D R R L A D P A T A R E W A G H Y R S V

23521 GCTCCGTGCGGTTCGTGCGCGACCCCGGCGTGCGCCTGTCCGCCCTCGGCACCCTGCTGTC 23580  
L R A V V A D P G V R L S A L G T L L S

23581 CCTGCCGCGACCGCCGTCGCCACGTCTTCGCGCGCCGGGAGATCGACGTCCGCGCGGT 23640  
L P R P P S A T S F G G R E I D V R R V

23641 CGAACGCGAGTTGGCGGGGCGCGACGGGATCACCGCCGCCCTGGTCGCGGTGGCGCCCCG 23700  
E R E L A G R D G I T A A L V A V A P R

23701 GCGCCTGGCCACCGGGTGC GCGTACGGGAACTGGTCGCCTACTGCGCCGTCGAGGGCAC 23760  
R L A T G L R V R E L V A Y C A V E G T

23761 GCCCGGTCCGAACGCGGCCACGACATCCGCGGCGCCTGCGGGAGCGCTGCCCGACGG 23820  
P R P N A A H D I R G R L R E R L P D G

23821 CTGGGTGCCGACCGTGTTCGTGAGCGCCCGCCGAGGAGATCCGGAAGGCCCTGGCCGC 23880  
W V P T V F V E R P P E E I R K A L A A

23881 CCGGGCGGCGGGCGGCGAACGGGCGGAGCCGCTGCCGCCGCCGAGGACTGCGTCCCGCT 23940  
R A A G G E R A E P L P P P E D C V P L

23941 TCCCGAGGAGGGGCCGCCCCCTCGGACCCGTCCGAGCGGCGGCTGGCCGCGCTCTGGGC 24000  
P E E G R P P S D P S E R R L A A L W A

24001 CGAGATCCTGGGCGCCCCGCGAAGAGCGTGACCGAGCCCTTCTCCGCGTCGGCGTCAC 24060  
E I L G A P P K S V T E P F F R V G V T

24061 CGATAAGGACGCCCTCCGCTTCTGCCCCGCGTGGCGGAGGACTTCGGCGTCACCGTGCC 24120  
D K D A L R F L A R V A E D F G V T V P

24121 CTTGCGCGACTTCTCAGCGCTCCCAACCTGCGTATGGTGAAGACAATTTGGCTGAGAA 24180  
F A D F L S A P N L R M V K D N L A E K

24181 ACGGAGGGTGTAACGCGCAATGAGTGAGTGGTAGGGTCGGAATCGAACCGCACTGATCGG 24240  
R R V \*

24241 CAATCTTTTCGGTCAGCTGTTCCGGATATTCGGGGCGCGTCGGCGCTCCCTCGACCAAG 24300

24301 GCGGTACGCGGATAAGCGTGCGCCGCCCCACGGCTGCGTCTCGACGCCTTCATCGGCGCG 24360

24361 TCGGACACTTCGCGGTGCCAGTCGGCACGCTCAGAGATCAGTGAATGCCTCGGTGTGCC 24420  
M P R C A (orf26)

24421 CGAGGTGCGCTCAGTACTGCTGTCCACACAACGCGCCAAGGGAGTTGGAACGTGATGGAG 24480  
R G A L S T A V H T T R Q G S W N V M E

24481 ACGGCGAATTCCGGCTATCGGGTCTCACCTCAGCAGCGGCATTATGGGCCATGCTGACC 24540  
T A N S G Y R V S P Q Q R H L W A M L T

24541 CGCGGGCGGGACGGCGGGGACGTCGCTTACCCAGTCCGCCGTGGTGGTTCGACCGTTCC 24600  
R G R D G G R R A F T Q S A V V V D R S

24601	CTGGACGCCGCACGTCTGCGCGCCGCGCTGGCCTCCGTGGTGGCCGCCACGAGCCGCTG	24660
	L D A A R L R A A L A S V V A A H E P L	
24661	CGGACGACCTTCACCGGTCTCGCGGGACGACCGCGCCGGTCCAGGTCGTCCATGACCCG	24720
	R T T F T G L A G R T A P V Q V V H D P	
24721	GACGAGCAGCCGCTGTCCGTCTGTCGACCTGCCGCCCTCGTGGCGCCGACGGCTCGGGCCCCG	24780
	D E Q P L S V V D L P P S C A D G S G P	
24781	GAACTGGACGAGCTCCGGCTCCGCGAACGCGCCGCCCTCGACCCGCGCGGGCGGCCCGTC	24840
	E L D E L R L R E R A A L D P R G G P V	
24841	TTCCGGGCGCCCTGGCGCGGGCGGCGAGGACCGGGCGGTGCTGGTGCTCACCGCGCAC	24900
	F R A A L A R A G E D R A V L V L T A H	
24901	GCCCTGGTTCGCGGACCGGTCTCCCTCCGGCTGCTGGCCGGGCGAGATCCTCGCGGCGTAC	24960
	A L V A D R L S L R L L A G Q I L A A Y	
24961	AGCGGGGAGACCGTGTCCCCGATGGCCCGCCGCTTGAGTACGCCGACTTCGCCGCC	25020
	S G E T V S P D G P P P L Q Y A D F A A	
25021	TGGCACCACGACCTGCTCACCGCCGAGGACGCCGCCCGACCGCGCGCACTGGGCGGCC	25080
	W H H D L L T A E D A A P D R A H W A A	
25081	CACACCGCCACCGCCGGCACCGGGCGCTCCCCGGCGTCTACGGCCCGGCGCCGCCCG	25140
	H T A T A G T G P L P G V V R P G A A P	
25141	GGTCCGTGGCGGGCGCGGGAGTGGGAACCTCCCCGCCGAACCTGGTGGCGGGGATCGACGGC	25200
	G P W R A R E W E L P A E L V A G I D G	
25201	GTCGCGGGAAGCTGTCCACCGATCCCGCCACCGTGTGTCACGCCGCTTCCGTATCGCG	25260
	V A G K L S T D P A T V L H A A F R I A	
25261	GTCTGGCGGCTCGCCGCGGAGCGGAACCTGCCCGTCCGCTCACTCGTGACGGCCGTTCC	25320
	V W R L A G E R N L P V A L T R D G R S	
25321	CACCCGAACCTCCGCACCGCGATCGGCGCCTTCGAGCGTGAGCTCCCGCTCGTCCACGAG	25380
	H P E L R T A I G A F E R E L P L V H E	
25381	ATCCGTACAGAGACGGCGTTTCGCGGAATACGCGCGCGCTCTGGACGCGCTCGTCGCCGAG	25440
	I R H E T A F A E Y A R A L D A L V A E	
25441	GGCGAGGAACCTCTCGACCATTCGACCCGGAACCTGCTCGGCAGCCTCGACGGCACCGCG	25500
	G E E L L D H C D P E L L G S L D G T A	
25501	GAAGGGCCCTGCTTACCTTACCCACCACCAGGCCGAAACACCGGTCCGGCGGGCGGGC	25560
	E G P C F T F T H H Q A E T P V R R A G	
25561	ATCACCTTTACCACCGTCCATCAGGATTCGGGTACGCCGATTCCCGTCCGCCTGACCGCC	25620
	I T F T T V H Q D S G T P I P V R L T A	
25621	CGACGCGACGGCGCCCGGTGCGCATGGAACCTGGGATACGACGAGGGCCGTATCGACGAG	25680
	R R D G A R L R M E L G Y D E G R I D E	
25681	ACGTTTCCCGAGAACGCCCGCGCTGCCTACCCGCATTCTCGAAGGCGTCTGCTCCGCC	25740
	T F P E N A A A C L T R I L E G V V S A	
25741	CCCAGGGGCCGGTTCGCGACATCCGCATGCTGTGCGACGAGACCGCACGGCTGCTCCGG	25800
	P E G P V G D I R M L S D E T A R L L R	
25801	GAAGCGGGGCTGGGCCCGCGTGGAACTTCCCGGCAAGGCGGTCCACGAACCTTTCGCC	25860
	E A G L G P R V E L P G K A V H E L F A	
25861	GAGCAGGCCGCGCGACCCCGGGCGGTGCGGGTCAGCGCGGGCGAGGACGCCCTCACG	25920
	E Q A A R T P G A V A V S A G E D A L T	
25921	TACCCGAACCTCGACGAGCGGTCCAACCGCTGGCACACCACCTGACCGGGCTCGGGGTG	25980
	Y A E L D E R S N R L A H H L T G L G V	
25981	ACACCCGGCCGGCACGTCTGGTCTCGGTGCGCCGCTCCGCCGAGCTGCTCGTGGGCTG	26040
	T P G R H V V V S V G R S A E L L V G L	
26041	CTCGGCGTCTCAAGGCGGGTGGCGCCTTCGTCCCCGTCGACGTGGGCTTCCCCGCAA	26100

L G V L K A G G A F V P V D V G F P R K  
 26101 CGGCTGGAGTTCGTGCTCCGGGAGACCGCCGCGCGGTCTCTGCTGACCGCCGACGTA 26160  
 R L E F V L R E T A A P V L L C T A D V  
 26161 CGGGACCGCATCGGCACTCGGACCCTCGACGACCGGGGTGACACCCGTGCGCTGGAC 26220  
 R D R I G T R T L D D A G V T P V A L D  
 26221 GCCGACCGGCGGCGCATCGCCGCACACCCCGCCGGCCCCACCGGCATCGCCACCACCCCC 26280  
 A D R R R I A A H P A G P T G I A T T P  
 26281 GACGCCCCCGCTACGTCGTCTACACCTCCGGCACCACCGGGAAGCCCAACGGCGTACGC 26340  
 D A P A Y V V Y T S G T T G K P N G V R  
 26341 GTCCCGCACCGGGGCTCAACCACTACCTACCTGGTGCACCGGCGCTACGGACTCGAC 26400  
 V P H R G L T N Y L T W C T G A Y G L D  
 26401 GGGGGCACCGGCACCCTCGTGACACCTCCATCAGCTTCGACCTCACCTCACCACCCTG 26460  
 G G T G T L V H T S I S F D L T L T T L  
 26461 TTCGCCCCCTGCTCGCCGGCGGGCAGGTGGTCATGCTCTCCGAGACCGCCGCGGTGACC 26520  
 F G P L L A G G Q V V M L S E T A G V T  
 26521 GGCCTGATCGCCGCGTGCCTCCCGGCGGACCTCACCTGGTCAAGCTGACCCCGACC 26580  
 G L I A A L R S R R D L T L V K L T P T  
 26581 CACCTCGACGTGCTCAACCACTGCTCACCCCCGACGAGCTGCGCGGCGCGGTCCGCACC 26640  
 H L D V V N Q L L T P D E L R G A V R T  
 26641 CTCGTGCTCGGCGGGGAGGCGGTGCGGGCGGAGAGCCTGGAGCCGTTCCGGGCCTCCGGG 26700  
 L V V G G E A V R A E S L E P F R A S G  
 26701 ACGCGGGTCTCAACGAGTACGGGCCCAGCGAGACGGTCTCGGCAGCGTCGCGCACGTC 26760  
 T R V V N E Y G P S E T V V G S V A H V  
 26761 GTCGACGCGCCACGCCCCGTACCGGCGCGGTGCCCATCGGCGGCGGATCGCCAACACC 26820  
 V D A A T P R T G P V P I G R P I A N T  
 26821 ACCGTCCACCTGCTCGACCAGCGGCGGCGGCGCCGTCCTCCGACGGCGTCTCGGCGAGCTG 26880  
 T V H L L D Q R R R P V P D G V V G E L  
 26881 TGGATCGGCGGCGCGGCTGTGCGCCGACGGCTACCTGGGGCGGCGGAACTCACCGGCGAG 26940  
 W I G G A G V A D G Y L G R P E L T G E  
 26941 CGCTTCTCTCCAGCGACTACCCGCGGACGGCGGCGCGGTCTACCGCACCGGCGACCTG 27000  
 R F L P S D Y P P D G G R V Y R T G D L  
 27001 GCCCGCGGCGCGCCGACGGCACCTGGAGTACCTCGGGCGCACCGACGCGCAGGTGAAG 27060  
 A R R R A D G T L E Y L G R T D A Q V K  
 27061 ATCCGCGGCGTCCGGGTGGAGCCCGCGGAGACCGAGGCGTCTCGCCTCCACCCCGGC 27120  
 I R G V R V E P A E T E A V L A S H P G  
 27121 GTCGGCCAGGCGCTGCTGGTTCGCGCGGCTGGACGAGACCCCGCGGTTCTGTCGCGCTC 27180  
 V G Q A V V V A R L D E D P G R S S P L  
 27181 GCCGCGAGCTGACGCTGACCGGCTACGTGTCCCGGCGCGGTGCCAGGCGCCCCCG 27240  
 A G E L T L T G Y V V P A R G A Q A P P  
 27241 CACGAGGAGCTCATCGCTACTGCCGGGAGCGGCTGCCCGAGCACTTCGTCCCGGCGCTC 27300  
 H E E L I A Y C R E R L P E H F V P A V  
 27301 CTCGTACCCCTCGACGCCCTGCCCGTACCGGCCACGGCAAGATCGACCGGCGTGGCTG 27360  
 L V T L D A L P V T G H G K I D R G A L  
 27361 CCCAAGCCGCACGCCCGGGCGGCGGACGGCGGCGGTACGTGCGCGCGCGCACCGCCACC 27420  
 P K P H A R A R D G A A Y V A P R T A T  
 27421 GAGGAGATCCTCGCGGCCACCGTCGCGAAGGTGCTGGGCGTCGAGCGCGTCGGCATCGAC 27480  
 E E I L A A T V A K V L G V E R V G I D  
 27481 GACAACTACTTCGTCTGGGCGGCGACTCCATCCGACGTCATGGTCGCCAGCCGGGCC 27540  
 D N Y F V L G G D S I R S V M V A S R A

27541	CAGGCCCGCGGGGTCGAGGTCACCGTGGCCGACCTGCACCGGCACCCACCGTCCGGGCC	27600
	Q A R G V E V T V A D L H R H P T V R A	
27601	TGCGCCGCGCACCTGGACGCGCGGAGGACCTGCCGCGGACGCCCGTCACCGAACCCCTTC	27660
	C A A H L D A R E D L P R T P V T E P F	
27661	GCGTGATCTCCGCCGAGGACCGGGCGCTGGTGCCGACGACGTGAGGACGCCTTCCCG	27720
	A L I S A E D R A L V P D D V E D A F P	
27721	CTGAACCTGCTCCAGGAAGGCATGATCTTCCACCGCGACTTCGCGGCGAAGTCGGCCGTC	27780
	L N L L Q E G M I F H R D F A A K S A V	
27781	TACCACGCCATCGCGTCCGTGCGGCTGCGCGCCCCGTTTCGACCTCGCCGTGCTGCGGATG	27840
	Y H A I A S V R L R A P F D L A V L R M	
27841	GTCGTGCGCCAGCTCGTCGAGCGGCACCCGATGCTGCGCACCTCCTTCGACATGAGCCGC	27900
	V V R Q L V E R H P M L R T S F D M S R	
27901	TTCAGCCGCGCGCTGCAACTGGTGACCGCGAGTTCGCCGATCCGCTGCACTACGAGGAC	27960
	F S R P L Q L V H R E F A D P L H Y E D	
27961	CTGCGCGGCAGGAGCGCCGAGGAGCAGGACGCCCGCTCGAGGAGTGGATCGAGCGGGAG	28020
	L R G R S A E E Q D A R V E E W I E R E	
28021	AAGGAACGCGGCTTCGAGCTGCACGAGTTCGCCGCTGATCCGTTTCATGGCGCAGCGCCTG	28080
	K E R G F E L H E F P L I R F M A Q R L	
28081	GAGGACGACGTCTTCCAGTTACCTACGGCTTCCACCACGAGATCGTGACGGCTGGAGC	28140
	E D D V F Q F T Y G F H H E I V D G W S	
28141	GAAGCCCTGATGATCACCGAGCTGTTTCAGCCACTACTTCTCGGTGATCTACGACGAGCCG	28200
	E A L M I T E L F S H Y F S V I Y D E P	
28201	ATCGCGATCAAGCCACCCACCGCCGGCATGCGCGACGCCGTCGCCCTGGAGCTGGAGGCC	28260
	I A I K P P T A G M R D A V A L E L E A	
28261	CTCGCGGACCGCCGCAACTACGAGTTCTGGGACTCCTACCTCGCCGACGCCACCCTGATG	28320
	L A D R R N Y E F W D S Y L A D A T L M	
28321	CGGCTGCCCAGGCCCGGCACCGGACCCCGGCCGACAAGGGCGACCGGGACATCACCCGC	28380
	R L P R P G T G P R A D K G D R D I T R	
28381	ATCGCCGTCCCCGTCCCCACCGAACTCTCCGACGGCCTCAAGCGGGTCGCCGCCACCCAC	28440
	I A V P V P T E L S D G L K R V A A T H	
28441	GCCGTCCCGCTGAAGACCGTGCTCCTGGCCGCGCACATGGTGGTGATGTCCCTCTACGGC	28500
	A V P L K T V L L A A H M V V M S L Y G	
28501	GGCCACGAGGACACCCTACCTACACCGTCACCAACGGCCGCCCGAGACCGCCGACGGC	28560
	G H E D T L T Y T V T N G R P E T A D G	
28561	AGCACCGCATCGGGCTGTTTCGTCAACAGCCTCGCGCTCCGCGTCCGGATGACCGGCGGC	28620
	S T A I G L F V N S L A L R V R M T G G	
28621	ACCTGGGCGGACCTGATCACCGCCACGCTGGAGTCCGAGCGCGCCTCGATGCCGTACCGG	28680
	T W A D L I T A T L E S E R A S M P Y R	
28681	CGGCTGCCGATGGCCGAACTCAAGCGCCACCAGGGCAACGAACCCCTGGCCGAGACGCTG	28740
	R L P M A E L K R H Q G N E P L A E T L	
28741	TTCTTCTTACCAACTACCACGTCTTCCACGTGCTCGACCGCTGGATCGACCGCGGCGTC	28800
	F F F T N Y H V F H V L D R W I D R G V	
28801	GGCCACGTGCGCAACGAGCTCTACGGCGAGTCCACCTTCCCCTTCTGCGGCATCTTCCGC	28860
	G H V A N E L Y G E S T F P F C G I F R	
28861	CTGAACCGGGAGACCGGCGAGCTGGAGGTCCGCATCGAGTACGACAGCCTGCAGTTCTCC	28920
	L N R E T G E L E V R I E Y D S L Q F S	
28921	GACGCCCTCATGGAGAGCGTCCGCGACAGCTACGCCCCGCTCCTCGCGGCCCTGGTCGCC	28980
	D A L M E S V R D S Y A R V L A A L V A	
28981	GACCCCGACGGGCGCTACGACCGGCACGAGTTCGCTCCGACCGCGACCGGGCCGCACTG	29040
	D P D G R Y D R H E F R S D R D R A A L	

29041	GCCGTCCTCACCCGCGGGCCCGAGGCGCGCGCGCCGACCGGTGCCTGCACGACCTGGTG A V L T R G P E A P A A D R C L H D L V	29100
29101	GCGGACCGGGCGGCGGACCGCCCCGACGCCCCGGCGCTCCAGCTGGACACCGACGTGCTC A D R A A D R P D A P A V Q L D T D V L	29160
29161	AGCTACGGCGAGCTCGACCGCCGCGCCAACCGGCTGGCCCACCACCTGCGTTTCGCTCGGC S Y G E L D R R A N R L A H H L R S L G	29220
29221	ATCGGCCCCGAGAGCGTCTGTCGGCGTCCTGGCCGAACGCTCCCTCGCCCAGATCATCGGC I G P E S V V G V L A E R S L A Q I I G	29280
29281	CTCCTCGCGGTCTCTCAAGGCGGGCGCGCCTACGTCCCGCTCGACCCGGCCCAGCCCCGAC L L A V L K A G A A Y V P L D P A Q P D	29340
29341	GAGCGCCTCGCCGCCGTATCGCCGGGAGCGGGCGCGCGCTCTCCACCGGCCCGGC E R L A A V I A G S G A A A V L H R P G	29400
29401	CTCGAAGGGCGGCTGCCCCGCGGGCGTCCGCGCGCTCCCCACCGACGCCGCCGACGGCAGC L E G R L P A G V R A L P T D A A D G S	29460
29461	ACCGCCACGCACGACCCCGGGCCCCACCGCCACGCCCCGCAACGCCGCTACGTGATGTAC T A T H D P G P T A T P R N A A Y V M Y	29520
29521	ACCTCCGGATCCACCGGAGAGCCCAAGGGCATCGTCGTGCAACACGCAACGTCGTGGCC T S G S T G E P K G I V V E H R N V V A	29580
29581	TCCCTCGCCGCCCCGCGCGCCCACTACGCGGCCGGACCCGGCCGGTTCCTGCTGCTGTCC S L A A R G A H Y A A G P G R F L L L S	29640
29641	TCCTTCGCCTTCGACAGCTCGGTGCGCGGCATCTTCTGGACGCTGACCCAGGGCGGCACC S F A F D S S V A G I F W T L T Q G G T	29700
29701	CTCGTCTGCCCCGGCGAGGGACAGCAACTCGACCCCGCGCGCTGGTGGAGACCATCGCC L V L P G E G Q Q L D P A A L V E T I A	29760
29761	CGGCAACGGCCCCACCCACACCCTCGCCATCCCCTCCCTGCTGGCGCCCGTCTTGACCAG R Q R P T H T L A I P S L L A P V L D Q	29820
29821	GCCGCCCCCGGCGACCTCGCCTCCCTGCGCACGGTGATCGCCGCGGGCGAGTCCTGTCCG A A P G D L A S L R T V I A A G E S C P	29880
29881	GCCGAACTGGCCGCGCCTGCCGGGACCTGCTGCCCGGGAGCACCTTCCACAACGAGTAC A E L A A A C R D L L P G S T F H N E Y	29940
29941	GGCCCCACCGAGACCACCGTGTGGAGCACCGTCTGGTCCCAGGAGAACGAGCACGACGGA G P T E T T V W S T V W S Q E N E H D G	30000
30001	CCCCACCTCCCCATCGGCCGGCCGGTTCGCGGGCACCTGGGTGACACCCCGCGACCACCGC P H L P I G R P V A G T W V H P R D H R	30060
30061	GGACGCACCGTCCCCCTCGCGCTCGCCGGCGAACTCTCCATCGGCGGCGCGCGCGTGGCC G R T V P L G V A G E L S I G G A G V A	30120
30121	CGCGGTACTCTCGGGCGCCCCCGGGACACCGCGCGCCGCTTCCGCCCCGACCCCGAGGCC R G Y L G R P R D T A A A F R P D P E A	30180
30181	ACGGCTCCCGGCGGCGCGCTACGCCACCGCGACCTCGGCGCTACTCTCCCGACGGC T A P G G R A Y A T G D L G R Y L P D G	30240
30241	AACCTGGAGTTCTCGGCCGCGCCGACCACGAGTCAAGATCCGCGGCTTCCGGGTGCGAG N L E F L G R A D H Q V K I R G F R V E	30300
30301	CTCGGCGAGATCGAGGCCGTCTCGACACCCACCCGGAGCTCCAGCGGACCATCGTCATG L G E I E A V L D T H P E L Q R T I V M	30360
30361	GCACGCGGCGACCACCCCGGCGACCAGGTGCTCGTCGCCTACGTCTCCCGCCCCCGGC A R G D H P G D Q V L V A Y V L P A P G	30420
30421	CGGCGGCCCCGAACCCGCGGACATCCAGGGGTACGTCCGCGACCGGTGCCCCGCTACATG R R P E P A D I Q G Y V R D R L P R Y M	30480
30481	GTGCCACCGCGGTGATCGTCTCGACGCGGTACCGCTGACCGCCCGCGCAAGGTGCAC G T G C C A C C G C G G T G A T C G T C T C G A C G C G G T A C C G C T G A C C G C C G C G G C A A G G T G C A C	30540



V P T A V I V L D A V P L T A A G K V D  
30541 CGGGCCTCGCTCCCCGCCCCAGCCACGCCAGCTCACCCGGGACCAGGAGTACGTTCGAG 30600  
R A S L P A P S H A Q L T R D Q E Y V E  
30601 CCCGGCACCGACACCGAGCGGGCGCTCGCCGCCATCTGGGCGGACGTCTCAAACCTGGAC 30660  
P G T D T E R A L A A I W A D V L K L D  
30661 CGGATCGGGGCCGGTGACCGCTTCTTCGACGTCTGGCGGCGAATCCCTGCGCGCGATGCAG 30720  
R I G A G D R F F D V G G E S L R A M Q  
30721 GCCACCGCGCGGCCAACAGATGTTCCGCACCCGCGTCTCCGTCCGCGCGCTCTTCGAG 30780  
A T A A A N K M F R T R V S V R R L F E  
30781 GCGCCTCCCTGCGGGAGTTCGCCCACGAGATCGACAAGGCCCGCTCGCGGGCGGCGGG 30840  
A P S L R E F A H E I D K A R L A G G G  
30841 ACCGGCCTACCGGCCCCGCGGCCGCCCGGCCACCGGAGGTGCCGCCGAATGACCCCGG 30900  
T G L T G P A A A P A T G G A A E \*  
M T P A (orf25)  
30901 CCGCCGACACCACCCACCGCTCTCGCCGGGCCAGCGCAGCATGTGGTTCTGCACCCGC 30960  
A D T T H P L S P A Q R S M W F L H R L  
30961 TCGCGCCCGAGGTGCCCGCTACAACATCTGCACCGCCATCGAGCTCACCGGCACACCGC 31020  
A P E V P A Y N I C T A I E L T G T P R  
31021 GCCCGGCGGCGCTGCGGGACGTGGTACGGCGGCTCGGCCGAGGCACGAGGCGCTGCGCA 31080  
P A A L R D V V R R L G R R H E A L R T  
31081 CGGTGTTCCCGTTCGGTGGGGAGACCCCCGCCAACGGGTCACCGACCGGGCGGCGCCCC 31140  
V F P S V G E T P R Q R V T D R A A P L  
31141 TGCGGACCGTGGACCTACCCACCTGACCCCGCGCCGCGGAGGCCGAGACCGCACGGA 31200  
R T V D L T H L T P A A A E A E T A R T  
31201 CGTACGGTGCGCCGCGCCCGGCCGTTCCGGCTCGACACCGGCCCCCTGGCGGAATGGA 31260  
L R C A A A R P F R L D T G P L A E W T  
31261 CCCTGCTGCGCGCGCCCCCGGCCACGCGTGTCTGCTCTCTCCGTCCACCACATCGTCT 31320  
L L R R A P G H A L L V L S V H H I V F  
31321 TCGACGGCGGCTCGTCCACGTGGTCTGCCGGAAGTGGAGGAGCGTACGGAGCGGCCC 31380  
D G G S L H V V C R E L E E A Y G A A L  
31381 TCGCCGGGCGCCCGGACCCCTCGGCACACCCGCGCGGGCTACGGACGGCAGTCCCGGA 31440  
A G R P D P L G T P A P G Y G R Q C R T  
31441 CGCGGGCGGCGGAACAGGACGAGGCCGGGCGGGAGTTCTGGCGCCGGAAGTGTCCGGCG 31500  
R A A E Q D E A G R E F W R R E L S G A  
31501 CGCCACCCCGCACGACCGTCTTCCGGGGCACCGGCCGGCCCGGACCGCCCGCCCGCG 31560  
P P R T T V F R G T G R P A G P P A R A  
31561 CCACCGTCCACTACGGCACCGACGATCCGGCCCCGACCGCGGACTTCTGCGCGAGCACG 31620  
T V H Y G T D D P A P T A D F C R E H A  
31621 CCGTACCCGGCTACGTGCTGCTGCTCGCGGCCCTCGCCTGCCTGGTCCGCCGGTACACCG 31680  
V T G Y V L L L A A L A C L V A R Y T G  
31681 GCCGACGGACGTGGTGATCGGCTCACCCGTTCGGAAGTGGCGAGGACCCCGAAGGGCTCG 31740  
R T D V V I G S P V G L R E D P E G L A  
31741 CCACCGTTCGGCCCGATGCTCAACCTGCTGCGCCTCGGCTGCACGGCGACCCCG 31800  
T V G P M L N L L P L R L R L H G D P G  
31801 GCTTCGGCGAGGTCTGGCCCGCACCCGGGAGACGCTGCTCGGCGCGCTGGAGCACCGCA 31860  
F G E V L A R T R E T L L G A L E H R T  
31861 CCACACCGTTTCGAGGACATCGTCGACGCGGTGGGCGCGGACCGGGACCCGGACGTCAGCC 31920  
T P F E D I V D A V G A D R D P D V S P  
31921 CCCTCTTCAGATCCTCTTCGCCCACGAACGCCCCCGGCCCCACCCGCGTTACCGGGCG 31980  
L F Q I L F A H E R P P A P P A L P G V

31981 TCCGTGCCCGCGTCGTACCCGTCCCCGTCCGGCCGCCAAGTACGAGCTCGCCGTCACCG 32040  
 R A R V V P V P A P A A K Y E L A V T A  
 32041 CCACCGAGACGCCCCGACGGGCTCCGGCTGATCGTCGAGGCGGAGCACGGACACGGGGAAC 32100  
 T E T P D G L R L I V E A E H G H G E P  
 32101 CGGCCGAACTCGCCGCTTCGCCCGCCACTTCGGCGTCTGCTGGCCGCCGGGGTCCGCG 32160  
 A E L A A F A R H F G V L L A A G V R A  
 32161 CGCCGACACACCGCTGAGCCGCTGCCGTGCTCACCAGCAGGAGCGGCGCCGGCTCA 32220  
 P D T P L S R L P L L T D E E R R R L T  
 32221 CCGACACCACGGCCCCCGCACCGCGCCGGAGGCCCCCTACCGCCCCCTGCACCGGCTGG 32280  
 D T T A P R T A P E A P Y R P L H R L V  
 32281 TCGAGGAGTCCGCCGCCCGCGGCCGACGCCCTGGCGGTCTGTCGGCGGCACGCGTCACC 32340  
 E E S A A R R P D A L A V V G G T R H L  
 32341 TCAGTACCGGGAGCTGAAC TGCCGCGCCAACCGCGTGCCGCTGGCTGCGCCGCGCTG 32400  
 S Y R E L N C R A N R R A A W L R R A G  
 32401 GCATCGGCACCGAGGACGTGGTTCGGCGTCCGGCTGGAACGCGGCCCGGAACCTCCTGCT 32460  
 I G T E D V V G V R L E R G P E L L V S  
 32461 CGCTCCTCGCCGCTCTCAAGGCCGCGCCGCTACCTGCCGTCGACCCGGCGCTGCCCCG 32520  
 L L A V L K A G A A Y L P V D P A L P A  
 32521 CCGAGCGGGTACGGCTGATGCTCGACGACGCCCGGGCCGCGTCTGCTCACCAGACCG 32580  
 E R V R L M L D D A R A A L L L T E T A  
 32581 CGTCTGGCACCCCGCCGGCCCCGGCCGGCACCCCGTGCACCACGTGGACGGACCGCCAC 32640  
 L G T P P A P A G T P V H H V D G P P P  
 32641 CGCCGACCCGGCCCGGGGACGACGCCGACACACCGGCCCGGACCTGCCACCAGCCTCG 32700  
 P T R P G D D A D H T G P D L P T S L A  
 32701 CCTACCTCCTCTACACCTCCGGGTTCGACGGGCCGGCCCAAGGCCGTGGCCCTCCAGCACG 32760  
 Y L L Y T S G S T G R P K A V A L Q H D  
 32761 ACAGCGCCCGCGCGTTCCTGCGCTGGGCGGGCCGCGCCTTCGACGGCGGGGAGCTGGCCG 32820  
 S A A A F L R W A G R A F D G G E L A A  
 32821 CCGTCTGGCCACCACCTCCGCCGGCTTCGACCTGTGCGTCTTCGAGCTGTTGCCCCC 32880  
 V L A T T S A G F D L S V F E L F A P L  
 32881 TGGCCACGGCGGCACCGTCTGCTCGCCGACAGCGCCCTGCACGTGCCCGCCCTGCCCT 32940  
 A H G G T V V L A D S A L H V P A L P W  
 32941 GGGCGCCCGCGCGACGCTCCTGAACACCGTGCCCTCCGCGGCCCGCCCTGCTGGACG 33000  
 A P A A T L L N T V P S A A A A L L D A  
 33001 CCGACGGCCTGCCGACGGTCTGACGGCCGTCAACCTGGCGGGCGAGCCCTGACCGCGG 33060  
 D G L P D G L T A V N L A G E P L T A E  
 33061 AGTGGTTCGCGCGGTGCACGCCCGCCTGCCGAAGGCCGCGTCCGCAACCTCTACGGCC 33120  
 L V A R L H A R L P K A A V R N L Y G P  
 33121 CCTCGGAGGCCACCACCTACGCCACCGCGGCCCTCGTGCCCGCGGGCGGCACCGAGGCGC 33180  
 S E A T T Y A T A A L V P A G G T E A P  
 33181 CGGCCATCGGCCGGGCGCTCGGCGCGGCCGCGTGTGGACCGCGACCGGCAGCGCC 33240  
 A I G R A L G A A R V W T A D D R Q R P  
 33241 CCCTCCCCGGCGCGGTCTGCTGGTGAACCTCATCGGCGGTACGGCCCCGGCCCGGGCT 33300  
 L P G A V V G E L L I G G T A P A R G Y  
 33301 ACCTCGGCCGGCGGGACCGACCGCCGACGCTTCGGGCCGATCCGACGGGACCGCCCG 33360  
 L G R P G P T A D A F R P D P T G P P G  
 33361 GCTCCCGGCTCTACCGCACCGGGGACCTGGCCGTACGCCGCCCGACGGCCGGTTCGTGT 33420  
 S R L Y R T G D L A V R R P D G R F V F  
 33421 TCCTCGGCCGCAAGGACGAGCAGATCAAAC TCCGCGGGGTGCGCATCGAACGGGCGAGG 33480

L G R K D E Q I K L R G V R I E P G E V  
 33481 TGGAAGCCGCTCTCCGCCAGTGC GCGCCGGTGC GCGCGGCCGCTCGTGCTCGCCGGA 33540  
 E A A L R Q C A P V A A A A V V L A G T  
 33541 CCACCGCGGAGAACCACCGCCTCGTCGGCTTCGTACCCCTTCGCCCCGGCGCCCGCTCG 33600  
 T A E N H R L V G F V T P S P G A R V D  
 33601 ACCCCGAGCGCACCCCTCGCCGCGCTGCGTTCGCGCCTGCCCGCCGCCCTCGTGCCCCCG 33660  
 P E R T L A A L R S R L P A A L V P A A  
 33661 CGCTGGTGGTGTGCGACGCCCTGCCGCTGACCGCCAACGGGAAGACCGACGGGGCCGCC 33720  
 L V V C D A L P L T A N G K T D R A A L  
 33721 TCGCCCGCGCGGCGCGGACACCGGCCGACACGGCGCTACGCCCCGCCCGCACCC 33780  
 A R R A R G H R P D H G A Y A P P R T R  
 33781 GCGTCGAGAAGGCGGTGCGCCGATCTGGCGCGAGGTGCTCGGGACCGAACGGGTGGGGA 33840  
 V E K A V A A I W R E V L G T E R V G I  
 33841 TCCACCAGGGGTTCTTCGACGCGGGCGGCACCTCCCTGTGCTGCTGCGCCTTCACCACC 33900  
 H Q G F F D A G G T S L S L L R L H H R  
 33901 GGCTGGTTCGCTCCGTCCATCCCGGCCTCCGGCTCGCGGACGTCTTCGGCTGCGGACCG 33960  
 L V A S V H P G L R L A D V F R L P T V  
 33961 TCGCCGCGCTCGCCGCGTTCGTGGACGGGACGAGGACGCGCGGAGACGGCCGTGCGCG 34020  
 A A L A A F V D G Q E D A R E T A V G D  
 34021 ACGCGGCCCTCCGGGCCGGCCGGCGCCGCGCGGTGGCCGCGCGCCGAGGAAAGGCG 34080  
 A A L R A G R R R A A V A A R R R K G G  
 34081 GCGGACGATGAGCCATGCCGACGCGGGCGACGGGCTCGACGCGGCTGACACGACTGACGC 34140  
 M S H A D A G D G L D A A D T T D A (orf24)  
 G R \*  
 34141 GGCCGACGGGATCGCCGTGATCTCGCTGGGCGGACGCTTCCCGGAGCGGACCGGGTGA 34200  
 A D G I A V I S L G G R F P G A D R V D  
 34201 CCGCTCTGGACGAACCTGCTCGACCGCGAGGACGCCATCAGCCACTTCACCGCCGACGA 34260  
 R L W T N L L D R E D A I S H F T A D E  
 34261 ACGCTCGCCCGGGGCCGACCCCGAACCTGGTGCGCCACCCCGGTTCTGTCGGCGCGGA 34320  
 R L A R G R D P E L V R H P R F V G A E  
 34321 AGGCGTCCTCGGCGACGTCTCCCTCTTCGACGCGGAGTTCTTCGGCTGCTCGCCGCGGA 34380  
 G V L G D V S L F D A E F F G C S P R E  
 34381 GGCCGAAGTCATGGACCCGACGACCGGCTCTGCCTGGAGGAGCGTGGCACGTCTTCGA 34440  
 A E V M D P Q H R L C L E E A W H V F D  
 34441 CACCGCCGGCTACGACCCGGCGGCGACGGGACCGCGGTGCGGGTGTCTCTCCGCGAG 34500  
 T A G Y D P A A T G T A V G V F L S A S  
 34501 CCTCAGCTCGTACCTGATCCGCAACGTCTGCCCGGCGGCGGCACAGCGCCTGCTCGG 34560  
 L S S Y L I R N V L P G G A A Q R L L G  
 34561 CGGCTTCCCGCTGCTGATCCACAACGACAAGGACTTTCTGGCCACCACCGTGTCCACAA 34620  
 G F P L L I H N D K D F L A T T V S H K  
 34621 ACTGGGCCTCACCGGGCCGAGTTACGCCGTGCGCTCGGCCTGCTCGTCCTCCCTCGTCGC 34680  
 L G L T G P S Y A V G S A C S S S L V A  
 34681 GGTGCACCTGGCCTGCCAGAGCCTGCTACCGAGGAATGCGACATGGCGCTGGCCGGCGG 34740  
 V H L A C Q S L L T E E C D M A L A G G  
 34741 GGTCTCGCTCCAAGTGCCGCGAGGGCCAGGGGTACGTGCACGCCGACGACGGCATCTACTC 34800  
 V S L Q V P Q G Q G Y V H A D D G I Y S  
 34801 ACCCGACGGGCGCTGCGCCCCCTTCGACGCCGGCGCGGCGGGCACGGTGGGCGGCAGCGG 34860  
 P D G R C A P F D A G A A G T V G G S G  
 34861 CGTGGGCCTCGTCCTGCTCAAGCGGCTCGCCGACGCCGTGCGCGACGGGGACCGGTCCA 34920  
 V G L V L L K R L A D A V R D G D R V H

34921 CGCGGTGATCCTCGGCTCGGCGGTGAACAACGACGGCGCCGACAAGGTCGGTTACACGGC 34980  
A V I L G S A V N N D G A D K V G Y T A

34981 GCCCGCGCTCACCGGCCAGAGCGCCGTCGTCCCGAGGCCCTGGCGGTGGCCGGGATCTC 35040  
P G V T G Q S A V V A E A L A V A G I S

35041 CGCCGCGACCGTCGGGCTCCTGGAGGCGCACGGCACCCGGCTGGGCGATCCCGT 35100  
A A T V G V L E A H G T G T R L G D P V

35101 CGAAGTGGCCGCGCTCACCCGGGCGTTCCGCGCCACACGGACCGCAGCGGCTTCTGCGC 35160  
E V A A L T R A F R A H T D R S G F C A

35161 GCTGGGCTCGGTGAAGGCCAACGTGGGCCACCTGGACGCGGCGGGCGTACCAGGGCT 35220  
L G S V K A N V G H L D A A A G V T G L

35221 GATCAAGGCCGTGCTGGCGGTCCGCGAGGGCGTCATCCCCGGCACCCCGCACTACCGTTC 35280  
I K A V L A V R E G V I P G T P H Y R S

35281 GCCCAACCCCGCCATCGACTTCGCCACCACCCCTTCTACGTACCGCCGACACCCCTCGC 35340  
P N P A I D F A T T P F Y V T A D T L A

35341 CTGGCCGGAGGCGGACCACCCCGCGGGCGGCGTCAGCTCCTTCGGCATCGGGGGCAC 35400  
W P E A D H P R R A G V S S F G I G G T

35401 CAACGCCACAGTGATCCTGGAACAGGCCCGCGCGCGCCCCCGCGCGGACCGGACCGC 35460  
N A H V I L E Q A P P A A P R A D R T A

35461 CGGGGTGCCCATGCCGTTGGTGGTGTCCGCCGACCCGCGAAGCACTGGCGGAGGCCGT 35520  
G V P M P L V V S A R T R E A L A E A V

35521 CCGGGACCTGGCGGCGTGGTCCGCCCGGAGCCGGGACCCGGCTCGCCGATCTCGCCGC 35580  
R D L A A W S A P E P G T R L A D L A A

35581 CACGTGGCCGGGCGCGGGCCTTCCCGTACCGCGCGCCGTCGTGTGCCACGACCTGCC 35640  
T L A G R R A F P Y R A A V V C H D L P

35641 CGAGGCCGCGCGCCTGCTGGGCGGCGCGCGGCGAGACCGCGCTCCCCGGCAGGGAGGC 35700  
E A A R L L G G A R G E T A L P G R E A

35701 CGTGTTCCTCTTCCCCGGGCAGGGCACCTCCCCCGGACACCGGGCGCGCCTGTACGC 35760  
V F L F P G Q G T L P P D T G R G L Y A

35761 GGACGTGCCGGCGTTCGCGCGCCACTTCGACGCCTGTGCCAAGGGTTCCGCCCGCTCGG 35820  
D V P A F R A H F D A C A E G F A P L G

35821 CACCGACCTCCACGCCGCGCTCGGGGCCCCGGCCGACACACAGGGCCGCGCAACCCGC 35880  
T D L H A A L G A P A D D T R A A Q P A

35881 CCTCTTCGCCGTGAGTACGCCCTCGCCCGCACCTGATGGACTGGGGTGTGCGCCCGGC 35940  
L F A V E Y A L A R T L M D W G V R P A

35941 CGCGATGCTCGGGCACAGCCTCGGCGAGTACGTCCGCGGACGCTGGCCGGGGTGTGTG 36000  
A M L G H S L G E Y V A A T L A G V L S

36001 CCTGCCGGACGCGCTGACGCTCGTCCGGGCCGGGCGGAAGCGCAGCACACCATGCCGCC 36060  
L P D A L T L V R A R A E A Q H T M P P

36061 CGGCCGATGCTCGCGGTCCCGCTCACGCCGACGACCTGCGCCGCTGCTGCCCCCGGA 36120  
G R M L A V P L T P D D L R P L L P P E

36121 GGTGGAGTTCAGCGCCTTCAACGCCCCCGGCGCTGCGTCTCGGCGGGCCCCCGAGCC 36180  
V E F S A F N A P G R C V V G G P P E P

36181 GGTGGCGGAGCTGCGCGCCCGCTGGCGCGGCGGAGTGCCGCGCCCGAAGTGGCCAC 36240  
V A E L R A R L A R R G V P A A E L A T

36241 CGCGCACGCCTTCCACTCGGCGGCCGTGAACCGCTGCTGGACGGCTTCCGGGGCGTGCT 36300  
A H A F H S A A V E P L L D G F R G V L

36301 GGAAGGCGTCCGACTGCGGCCGCCCCGGCTGCGGTACGTGTCTCCCTACCGGCGACTG 36360  
E G V R L R P P R L R Y V S S L T G D W

36361 GGCCGACGCCGCGGTACCAACCCCGCGTACTGGCTCGCCACCTGCGCCGGCCCGTCCG 36420

A D A A V T T P A Y W L A H L R R P V R  
 CTTCGCCGACGGCCTGCGGCCTGCCTGGACCTCGGCCCGTCCGCCCTGGTCGAGACCGG F A D G L R R C L D L G P V A L V E T G  
 GCCGCGGGCCGGACTGACCGGCCTGGCCCGCCGCGCCGCGGGCCCCGGCGAGCCCCCTTA P R A G L T G L A R R A A G P G E P P Y  
 CACCGTCCGCTGCCTGGCCGCCCCGACGAGGCGGCTTCGCTGACCCACGCGGTCCGCGT T V R C L A A P D E A A S L T H A V A V  
 ACTCTGGCGCTCGGGCTGCGCCGTCGACTGGACGGCGTTCCACCGCCCCGGGCGCCCCCG L W R S G C A V D W T A F H R P G R P R  
 CCGCACCCACGTGCCCCGCTACCCCTTCCACGGGTACGGCACTGGATCGACGCGCCGGA R T T V P G Y P F Q R V R H W I D A P D  
 CGAGTCCGAACCCACGGACCTCGCCACCGCCCTGCGCGCGGAGTTGCGGACGGACGGCGA E S E P T D L A T A L R A E L R T D G D  
 TCCGCGCTCGCCGTCGATCAGCGGCCCGACTGCGCACGGGGCTGAACCGGCTGTGCGC P P L A V D Q R P G L R T G L N R L C A  
 CGCCCTGGCCCGGACTACCTGGCCACCGCGCTCGAAGCGAGCGGGTCTGCCCCGATT A L A R D Y L A T G V E A S G V L P G F  
 CCACCGCTTCTGGACTACCTGCGCACCTTGGCCGCTCCGACCGGCCGCGGACGACGC H R F L D Y L R T L A A S A P A A D D A  
 GGGGACGATCGCCCGGAGATCACC CGGCCCCACCCGTCCTTCTCCGGGCTCGTCGACCT G T I A A E I T A A H P S F S G L V D L  
 GCTCCGGCACTGCGCCAGGGCTATCCGCGCGCCCTGTCCACCCCGGAGCCGCACTGGA L R H C A Q G Y P R A L S T P G A A L D  
 CGTCTCTATCCGGCCGGCAGCGGCGACCTCCTGCGCGCACCTGGGCGAGGGCACCGC V L Y P A G S G D L L R R T L G E G T A  
 CGACCACCGCGCCACCGCGCCTCACCCGCTGGCCGGCTCCCTGCTCGACCGGCTCGC D H R A T G R L T R L A G S L L D R L A  
 GGCCGACCGCGAACCCGGCCCGCCGCTGCGCGTCTGGAGGCCGAGCGGGCGCGGGCAG A D R E P G R P L R V L E A G A G A G S  
 CCTCACCCAGGCCCTGGTCACCCGGGCCCCCGGCCGGCTCGACTACCACGCCACCGACAT L T Q A L V T R A P G R L D Y H A T D I  
 CTCCCGGCACTTCGTGACCGCACTCGGCCGGAGGCCCGCCGCGCGGCTGGACTTCGT S R H F V T A L G R E A A R R G L D F V  
 CCGCGCACGCGTCTCGACATCGCCCGGACCCAGGCGAACAGGGCTTCGCCGGCGAGCG R A R V L D I A R D P G E Q G F A G E R  
 GTTCGACGTCTGCGGCCTCGACGTGGTCCACGCCACCCCGACCTGCGCACCCACGCT F D V V C G L D V V H A T P D L R T T L  
 CGGCCATCTGCGCTCCCTGATGGCACCGGACGGCACCTCGCGCTGATCGAGACCACCGC G H L R S L M A P D G T L A L I E T T A  
 CGACGACCCCTGGCTGACGATGATCTGGGGCTGACGGACGGCTGGTGGCACCACACCGA D D P W L T M I W G L T D G W W H H T D  
 CCGGCGCACCCACGGCCCGTCTGCTCGACGCCCGGCTGGCGCGCCCTCCTGGCCGGCGA R R T H G P L L D A A G W R A L L A G E  
 GGA CTTCGCCACGGCCGATGTGATCGTGCCGCCCGACGGCCCCAGGACGCGGCCCTGCT D F A T A D V I V P P D G P Q D A A L L  
 GCTCGCCCGGACAGCCCCCGGCCGGCGGCCCGCACCGTCCGTCGGCAAGCGGGACGT L A R Q T P R P A A A A P S V G K R D V  
 CGGCACGTGGTGCTACGCCCGGGCTGGCGGCACGCCCGCGCCCGGACCCCGCCCCGCT G T W C Y A R G W R H A A P A D P A P L

GACGGGCGGCTGCCTGCTGCTGGGCGACGGGGACACGGCGAAGGCCGTCGCGAGCCGGCT 37920  
T G G C L L L G D G D T A K A V A S R L  
GGAGGCCCTCGGCGTGCCCGTCACCACCGTCGGCGGGCGGCGGACCGCGGGGCCCCGAGCG 37980  
E A L G V P V T T V G G G R P P G P E R  
GTACCGGGAACCTCGTCGGCCCCGCCACCCGCTGGCCGTCGACCTGTGGCCGCTGCGCGA 38040  
Y R E L V G P A T R L A V D L W P L R D  
CGCGTCCCACCGCGGCGCGCGCGCGGCGCGCGCGGCGTACGGACCGCCCAGGACGCGCGC 38100  
A S H R G R A A G A A G V R T A Q D A A  
GCTGCACAACCTGCTCCACCTCGCCCCGGGCCTTCGGCGCGCTGGAGGAGCGCCACCCCGC 38160  
L H N L L H L A R A F G A L E E R H P A  
CCGCGTCGTGACCGTGACCACCGGTGCCACGACGTGCTCGGCGACGACCTCGCCCCACC 38220  
R V V T V T T G A H D V L G D D L A H P  
CGAGCACGCCACCGTCCCGCGCGCGGCAAGGTGATCCCCCGGGAGTACCGTGGATCGC 38280  
E H A T V P A A A K V I P R E Y P W I A  
CTGCACCGCCCTGGACGTGGAGCGGGCCCTGGACGCCGAGCGGCTGGCGGACCTGATCGT 38340  
C T A L D V E P G L D A E R L A D L I V  
CCGGGAACCTCGGCGCGGCGCGCGAGACCACCGTCACCGCCTGCCGCGGCGGACGCGCCTT 38400  
R E L G A A R E T T V T A C R G R R R F  
CACCCCTGCCCCGTCCGGCAGCCCCCTCCCCCGCGCACCGGAACGCCCGGGCGGTCCGGCC 38460  
T P C P V R Q P L P A A P E R P A V R P  
CGGCGCGCTCTACCTCGTCTGCGGCGGGCCTCGGCGGCATCGGCCTCCACCTCGCCGAGTA 38520  
G G V Y L V C G G L G G I G L H L A E Y  
CCTGGGCGCGCGCGCACCCACCGTCGTCCTACCCACCGGCGGCCCTTTCCCGCCCCCGG 38580  
L G R A R T T V V L T H R R P F P A P G  
CGCGTGGGACGGGCTGCCCGCGGGACACCCGAGGCGGCGCTCGTCCGGCGGCTGCGCTC 38640  
A W D G L P A G H P E A A V V R R L R S  
CCTCGCCGCCACCGGCGCCACGGTCGTCGTCGCGCGGGCGGACCTCACCGACCACGACGC 38700  
L A A T G A T V V V R R A D L T D H D A  
GATGCGCGCCCTCGCGGACGAGGTGGAACAGGCCCACGGCCCCGTCCGGGGGGTGGTGCA 38760  
M R A L A D E V E Q A H G P V R G V V H  
CGCGGCGGGGTGCCCCGACACCGCGGCATGATCCAGCGTCGCGACCGAGCCGGCACGGA 38820  
A A G V P D T A G M I Q R R D R A G T D  
CGCCGCCCTCGCCGCCAACTGACCGGCACCCCTCGTCCTGGACGAGGTGTTCCGCCACCG 38880  
A A L A A K L T G T L V L D E V F A H R  
CGACCTCGACTTCCTCGTCCTGTGCTCCTCGATCGGCACCGTGTGCACAAGCTGAAGTT 38940  
D L D F L V L C S S I G T V L H K L K F  
CGGCGAGGTGCGCTACGTGGCGGGCAACGAGTTCCTCGACGCCTATGCCGCCACCGCGC 39000  
G E V G Y V A G N E F L D A Y A A H R A  
GGCCCGCGCGCGCGCAGAACCCCTGTGATCGCCTGGACCGACTGGCGGGAGTCGGGCAT 39060  
A R R P G R T L S I A W T D W R E S G M  
GTGGCGCGCGCGCCAGCGCGCTGTGACCGAGCGCTACGGCACCGGCGCGGACCTGCCCGT 39120  
W A A A Q R R L T E R Y G T G A D L P V  
ACCGCCCGGGGGCGACCTGCTCGGCGCGATCAGCCCCGAGGAGGGCGTCGACGTCTTCGC 39180  
P P G G D L L G A I S P E E G V D V F A  
CCGGTGTCTCGCGCGGACACCGGCGCGAACGTATCGTGTGCGCCAGGACCTCGACGA 39240  
R L L A A D T G P N V I V S A Q D L D E  
ACTCCTCGCGCGGACGCGCGGTACACCACCGACGACCTCGCGGCCCTCGGCGACCT 39300  
L L A R H A A Y T T D D H L A A L G D L  
GAGGATCGCCCGCGCGGACCGCTCCGCGCGCGCGCGGTACGCGGCCCCCACAC 39360  
R I A A A R D R S A P A A P Y A A P H T

39361 GCCCGCCCAGCGGCGGATCGCCGGCTGGTACCGCGACCTGCTCGGCGTCGAACACGTCGG 39420  
 P A Q R R I A G W Y R D L L G V E H V G  
 39421 CCTCGACGACGACTTCTTCGCGCTCGGCGGGGACTCGCTGCTCGCCCTGCGCCTGCTGTC 39480  
 L D D D F F A L G G D S L L A L R L L S  
 39481 GCAGTGCGGGACGCCTACGGGGTGGAGATCTCCGTGCGCCGCATGTTGACGAGCCAC 39540  
 Q L R D A Y G V E I S V A R M F D E P T  
 39541 GGTGGCGGCGCTGGCCGCCGCCACCGGCCCGCCGCGGAAGAGACGCCCGCCAGGAAGA 39600  
 V A A L A A A T G P P P E E T P G Q E E  
 39601 GGTGGTGCTGTGACCACGCCCGCATCACCGACCTGCTCACCGAGCTCCGCGGCCGCGCAG 39660  
 V V L \*  
 M T T P R I T D L L T E L R G R Q (orf23)  
 39661 GTGACCCTCACGGCCGACGGGGACCGGCTGCACTGCCGCGCGCCCGGGGCGCGCTCACC 39720  
 V T L T A D G D R L H C R A P R G A L T  
 39721 GACGAGCTCCTCGCCACCATCCGCGCCCGCCGCGACGAACCTCCTCGCCACCTGCGCGCC 39780  
 D E L L A T I R A R R D E L L A H L R A  
 39781 GACCGCCGCATCCCGCGCCACGACGGGCCCGCGCCGCTGTCTTTCGCCCAGGAACGGCTC 39840  
 D R R I P R H D G P A P L S F A Q E R L  
 39841 TGGCTCTCCACCAGTTCCACCCGCACGACAGCGCCTACAACATCCCCCTGCACATCGCC 39900  
 W L L H Q F H P H D S A Y N I P L H I A  
 39901 CTGCGCGGGCCCCTGAACCCGGCCGCCCTGCGCGCCGCCCTGGCCGAGGTGGTACGGCGG 39960  
 L R G P L N P A A L R A A L A E V V R R  
 39961 CACGACGTCCTGCGCACCCGGTACGCCATCAGCCGCGGCCTGCCCCGGCCCGTCGTCGAA 40020  
 H D V L R T R Y A I S R G L P R P V V E  
 40021 CCGGCCCACACGCGCGCGCTGCCCCTGACCGACCTGACCGGGCTCCCCGCACACCACCGG 40080  
 P A H T P P L P L T D L T G L P A H H R  
 40081 GACGCCGAACCTCGCCCGGCTGGCCGCCAGGAGGCCAGGCGGCCCTTCGACCTCGCCCAG 40140  
 D A E L A R L A A Q E A R R P F D L A Q  
 40141 GGCCCGGTGCTGCGGGCCCGGCTCCTCCGAACGGCCCCCGAGGAGCACCGGCTGCTGCTG 40200  
 G P V L R A R L L R T A P E E H R L L L  
 40201 ACCCGCCATCACATCGCCAGCGACGGCTGGTTCGCTCGACATCCTGCTCCGCGAACTGGGC 40260  
 T R H H I A S D G W S L D I L L R E L G  
 40261 ACGTTCTACCGGGCAGGGCGGGACGGCACACCCCGCGGCTTCGACGCCCTGCGGCTGCGG 40320  
 T F Y R A G R D G T P A G L D A L P L R  
 40321 TACGCCGACTTCGCCCGGTACCAGCGCGAACAGGCCGAACGGCCGGAGACGGCCGAGCGG 40380  
 Y A D F A A Y Q R E Q A E R P E T A E R  
 40381 TCGACCCGCTGGGCACGGCACCTGAGGGGCGCCCCCGCGACACTCGACGTCCTCGGGCCC 40440  
 S T R W A R H L R G A P A T L D V L G P  
 40441 CCGCCCCCGGAACCTCCACGCGCCGGCCGGCACCGTACGGACGGACCTTCCCGCCGCC 40500  
 P P A E P S H A P A G T V R T D L P A A  
 40501 CTCGTCACCGGCCTGCGGCAGCTGGGCGGCCGGGCCCGCACACGCTCTTCCCGCTCCTG 40560  
 L V T G L R Q L G G R A R T T L F P L L  
 40561 CTGAGCGCCTTCGGCCTCGCCCTGGCCGGGCCCGCCCGCCCGTACGACGTCATGGTCGGC 40620  
 L S A F G L A L A G P P G P Y D V M V G  
 40621 ATCCCCGTGCGCCGGCCGGCCGCGCACCGAACTGGAGCCGCTCATCGGCTGCTTCGCGACC 40680  
 I P V A G R P R T E L E P L I G C F A T  
 40681 ATCGCGCCGATGCGGCTGACGAGCGACGGGACCGAGCCGCTGACCCGGCTCGCCGCCCGC 40740  
 I A P M R L T S D G T E P L T R L A A R  
 40741 GCCCAGCAGCAGTCCAGGACGCGCTGGACGGACCCGACGTCCCCTTCGAGCGGCTCGTG 40800  
 A Q Q H V Q D A L D G P D V P F E R L V





42301 CTGCCCCACGGCGGCCTGGACTACGCGGGCCGCTCCGACGCACAGGTCAAGGTCCGCGGC 42360  
L P D G G L D Y A G R S D A Q V K V R G

42361 TACCGCGTCGAGCCCCGCGAGACCGAAGCCGCGCGCTGACCCATCCCGCCGTGCGCCAC 42420  
Y R V E P A E T E A A A L T H P A V R H

42421 TGCGTGGTCGTGCCACGCGGCGACGGCGACCGGCCATCTCGCGGCGTACGTCTGCGCC 42480  
C V V V P R G D G D R R H L A A Y V V A

42481 GACACCCGCGCCTGCGACGGGCCCGGGCTCCGACCCACCTGGCCGAGCGGTGCCCGC 42540  
D T R A C D G P G L R T H L A E R L P R

42541 CACCTGGTGCCGGCCTCGGTGGTCTTCTGAAGCGGATCCCGCTGACCCGCAACGGCAAG 42600  
H L V P A S V V F L K R I P L T R N G K

42601 CTCGACGTGGCGGCCTTGCCCGACCCGGCGCCACCGCGCACCCGCCCGGAACGCCCG 42660  
L D V A A L P D P A A H R A P A R E R P

42661 CGCACCCGCGACCGAACGGACCTCACCCGGGTGCTCGCGGCCCTCTGAAGGCGCCACCG 42720  
R T A T E R T L T R L L A A L L K A P P

42721 GAGACCATCGGGACGCACGACAACCTCTTCGACCTGGGCGGCGACTCCCTGACGGTCACC 42780  
E T I G T H D N L F D L G G D S L T V T

42781 CAGTTCCTACTCCCGGGTGGTGGAGGAGTTCGCGGTGACCTCCCGGTGCGCCGGGTCTAC 42840  
Q F H S R V V E E F A V D L P V R R V Y

42841 CAGGCCCTCGACATCGCGACGCTCGCCGTGACCGTGGACGACTTCCGGCGCCGCGCCGAA 42900  
Q A L D I A T L A V T V D D F R R R A E

42901 CGCACCCGCGGTACTGCGCGCCCTCGCGGCGGCGGAGGCGATGGAACCCGGCGGTACGGCG 42960  
R T A V L R A L A A A E A M E P G G T A

42961 GGGGAGTCCGGCGGTAATCCGAGGAGTCCGCCGCTACGGCGCGGGGGCCCGCCGTGCGG 43020  
G E S G G N P E E S A A T A R G P A V A

43021 GCGAACGAACCCGGCGCTGCGGCGCGTGAGTCCGGCGCCGCGCCGGTGGAGCCCGCCGTC 43080  
A N E P G A A A R E S G A A P V E P A V

43081 GCAGTACAGGAGTCCGCCGTACGAAGGGGGAGCCCGGCACCGCAGCGAATGAACCTCGGC 43140  
A V Q E S A A T K G E P G T A A N E L G

43141 GCTGAGGCACGGGAGCCCGGCACCGCAGCGCAGGAACCCGGCACCGACCCCGGCCACCC 43200  
A E A R E P G T A A Q E P G T D P R P P

43201 GCCGCCACACCGCAGGACCCCGCACACACCGCAGGAAGGACAGCCGTGCCCGCGTCCC 43260  
A A T P Q D P R T T P Q E G Q P C P R P

43261 GAATGAGCCGGCCGGCCGGCATCGTCGACATCGCGCGCCGTACGCGGAGCGCACCCCG 43320  
M S R P A G I V D I A R R H A E R T P A (orf22)  
E \*

43321 CCCGTCCCCTGACGCGTTCCTGCCCCACGGCGAGACGGAGAGCGTCCGCTTCTCCTTCG 43380  
R P A Y A F L P D G E T E S V R F S F A

43381 CCGACATCGACCGGCGGGCCCGCGCCGTGGCCCGCTCCTCCAGGACCGCGGCCTGGCCG 43440  
D I D R R A R A V A A V L Q D R G L A G

43441 GGGAGCGGGTCTGCTGCGCTATCCCTCCGGGCCCCGAGTACGTCCAGGCGTTCCTGGGCT 43500  
E R V L V A Y P S G P E Y V Q A F L G C

43501 GCCTGTACGCGGGCGTGGTCCCGTCCCCTGCGACGAGCCGCGTCCGGCCCCGAGCGCGG 43560  
L Y A G V V A V P C D E P R S G P S A E

43561 AACGGCTCGCCGGGATCCGCGCCGACGCCCGCCCCGCCCTGGCCCTGACCGCCGGCGCCC 43620  
R L A G I R A D A R P A L A L T A G A P

43621 CCGAGGCCGGGCTCGCCGCGCTGGCCACCTGGACGTGGCCGGCGTCCCCGACTCCGCCG 43680  
E A G L A G L A T L D V A G V P D S A A

43681 CCGGGGCTGGACCGACCCCGTCCGCGGACCGGACGCCCTGGCCCTTCTCCAGTACACCT 43740  
G A W T D P V A G P D A L A F L Q Y T S

43741 CCGGATCGACCCGCCGCCCGCGGCGTCATGGTCGGCCACGGCAATCTGCTGGCCAACG 43800  
G S T R R P R G V M V G H G N L L A N E

43801 AGCGTGTCATCGCGCGCCTGCGGCCACGACCGGACTCCACCTTCGTGGGATGGGCGC 43860  
R C I A A A C G H D R D S T F V G W A P

43861 CGTTCTTCCACGACATGGGCTGGTCGCCAACCTCCTCCAGCCCCTCTACCTCGGGTCCC 43920  
F F H D M G L V A N L L Q P L Y L G S L

43921 TGTCCGTGCTGATGCCGCCGATGGCCTTCCTCCAGCGCCCGCCCGCTGGCTGCGGGCCG 43980  
S V L M P P M A F L Q R P A R W L R A V

43981 TCTCCGCTACCGGGCGCACACCAGCGGCGGCCCAACTTCGCCTACGACCTGTGTGTCG 44040  
S R Y R A H T S G G P N F A Y D L C V D

44041 ACCGGGTCGGCGAGGACGAGCGGGCCGACTGGACCTGTGCGGGCTGGAAGGTCGCCTACA 44100  
R V G E D E R A G L D L S G W K V A Y N

44101 ACGGCGCGGAACCTGTACGGGCCGACACCCTGCGACGTTTACCGACCGCTTCGCCCCC 44160  
G A E P V R A D T L R R F T D R F A P H

44161 ACGGCTTCACCCCCGCGCGCACTTCCCGACCTACGGGCTCGCCGAGGCGACCTGCTCG 44220  
G F T P G A H F P T Y G L A E A T L L V

44221 TCGCCACCGGCCCCAAGGAGTGCCGCCCCGCAACCCTGACCGCCGACCGCGCCGCCCTGC 44280  
A T G P K G V P P R T L T A D R A A L R

44281 GCGCCGCGCGGCTCCGGCCCCGCGGGCCGCGAGGCGGCCTGGAAGTGGTCGGCAACG 44340  
A G R L R P A G P G E A G L E L V G N G

44341 GCACCGCGGCCTCGACACCACCTCCGGATCGTCGACCCCGCGACCGCGCGGGAGTGCC 44400  
T A G L D T T L R I V D P A T A R E C P

44401 CGCCCGGAGAGGTGCGCGAGGTCTGGGTGCGCGGCCCGGGCGTGGCACGCGGCTACTTCG 44460  
P G E V G E V W V R G P G V A R G Y F G

44461 GCCGCCCCGCGCGAGTCCGCGCCGCTGCTCGCGCCCGCCTGCCCGGCGGCGAAGGACCGT 44520  
R P R E S A P L L A A R L P G G E G P Y

44521 ACCTGCGGACCGGGGACCTGGGCGCCCTGCACGACGGGGAACCTCTCCTACCGGACGCC 44580  
L R T G D L G A L H D G E L F L T G R H

44581 ACAAGGACCTCATCGTCATCCGCGGCCAGAACCACCACCCGACGACCTCGAACGGACCG 44640  
K D L I V I R G Q N H H P H D L E R T A

44641 CCGAGCAGGCCCCACCGCGCTCCGCCCCGACCTGCGCCGCGCGTTCGCGGTGCCCGGG 44700  
E Q A H P A L R P T C A A A F A V P G D

44701 ACGGCGCGGAGCGGCTCGTGCTCGTCTGCGAACTCACCTCCTACCGCGCCGTCGACCCGG 44760  
G A E R L V L V C E L T S Y R A V D P A

44761 CCGCCGTGCGCGAGGCGCTCCGGGCCGCGCTCGCCGCGCGGCACGGCGTCCGCCGCACA 44820  
A V A E A V R A A L A A R H G V A P H T

44821 CGCTGGTGGTGCTGCGCCGCGGCGGCATCCCCAAGACCACCAGCGGAAAGGTGCGGCGCG 44880  
L V V L R R G G I P K T T S G K V R R G

44881 GCCACTGCCGACGGCCTACCTCGACGGAACGCTCCCCGTTACACGGCCGTCCGCCTCC 44940  
H C R T A Y L D G T L P V H T A V R L P

44941 CGGCGGGGGAGGAGGGCACCGAGGCCCTTCCCTGACCACGGACCCCGGTGCGCTGGCCA 45000  
A G E E G T E A L P L T T D P G R L A T

45001 CGGCGTGCAGACCTGGCCCGCCCCACGCGGCGCTGGCCGGGCCCCCTCCCCGGCACCG 45060  
A L R D L A A A H A G L A G P L P G T D

45061 ACGAGCCGGTGAGCGCCCTCGGCCTGGACTCGCTCGCCTCCCTGCGGCTCCACCACCACG 45120  
E P V S A L G L D S L A S L R L H H H V

45121 TCCAGTCCGCTACGGCGTGACCTGCCCCGTACCGCCCTGCTCGGCGACACCACTTACC 45180  
Q S A Y G V T L P V T A L L G D T T Y R

45181 GCCGGCTCGCGGAGCTGACGCTCGCCGCCCCCGCCCGGCGCCCGAGGGGCAAG 45240  
R L A E L T L A A P R P A R A P E G Q V

45241	TCACCGGCGTCTGGCGGCCGTTGACGCACGGGCAGCGCGCCCTGTGGTACGAACAGGCGC	45300
	T G V W R P L T H G Q R A L W Y E Q A L	
45301	TCGCCCCGCACGCGGCCGCTACCACCTCGTCCGCGCGCTGGCCCTCCGCGGCCCGCTCG	45360
	A P H A A A Y H L V R A L A L R G P V D	
45361	ACGAGGAGGCCCTCGCCGAGGCGGTCCGCCGCGTCCGCCGCCACCCCGCCCTGCGGA	45420
	E E A L A E A V R R V V R R H P A L R T	
45421	CCCGCTTCGCGCTCCGCGACGGCGAACCGGCGCGCCGACCGAGCCGTACGGACCGGAGC	45480
	R F A L R D G E P A R R T E P Y G P E L	
45481	TGGACGTACGCGACGCCACCGGCTGCCGGCGGACCGGCTCCGCGAACACCTGGCCGCGG	45540
	D V R D A T G L P A D R L R E H L A A A	
45541	CGGGCGACCGCCCTTCGACCTGGCCGCGCGACAGGCCCGTGAGGCTGACGCTCTACC	45600
	G D R P F D L A A G D R P V R L T L Y R	
45601	GCACGGACGGCGGCCACATCCTGCTGCTGGTCGCCCCACCACCTGGTCGCGACTTCTGGT	45660
	T D G G H I L L L V A H H L V A D F W S	
45661	CCCTCGTCGTCCTCTGGGCGACCTCGCCCGGGCCACGCGGGCGAGGACCTGCCGCCCCG	45720
	L V V L L G D L A R A H A G E D L P P A	
45721	CGCCGGAGGGGGACCCCGCGACGAGGCGACGGACGCGGACCGGACGTACTGGCGGCACC	45780
	P E G D P G D E A T D A D R T Y W R H R	
45781	GGCTCGCCGACGCGCCACCCGCCCTCGACCTGCCCCACCGACCTCCCCACCCCGCCGAGC	45840
	L A D A P P A L D L P T D L P H P A E R	
45841	GCGGCTTCGCGGCGCCACCCACGCCTTCGGCTGCCCGGACCTCACCGCCCGGCTGA	45900
	G F A G A T H A F R L P P D L T A R L T	
45901	CCGCCCTCTCCCGGGAACGGCACTGCACCCTCTTACCACCCTCCTCGCCGCCACCAGC	45960
	A L S R E R H C T L F T T L L A A H Q L	
45961	TACTGCTCCACCGCCTGACCGGGCAGGACGACCTCGTCGTGGGCACCCTCCTCGCCCGCC	46020
	L L H R L T G Q D D L V V G T L L A R R	
46021	GCGACACCGCCGAAGCGGCGCGCCGTCGGTACCTGGTCAACCCGCTGCCGCTGCGCT	46080
	D T A E A A G A V G Y L V N P L P L R S	
46081	CCGTACGGGAGCGGGGAGACCTTCACGGAACGCTGCGCGCACCCGGCGGACCGTGC	46140
	V R E P G E T F T E L L R R T R R T V L	
46141	TGGACGCGGTTCGCGCACGGCCGCCACCCCTCGGGCGGCTCGTCTCCCGTCTCGCCCCCG	46200
	D A V A H G R H P F G P L V S R L A P A	
46201	CGCGCACGCCCGCGCGCGCCGCTCCTGCAGAGCCTGTTCTGTGCTCCAGCGCGAGTACG	46260
	R T P G R A P L L Q S L F V L Q R E Y G	
46261	GCGACGAGGCGGACGGGTACCGCGCGCTCGCCCTGGGCGTCCGGCGGCGGCTGCGCGTCG	46320
	D E A D G Y R A L A L G V G G R L R V G	
46321	GCGGACTCGACCTGGAGGCACTCGCGTTGCCGCGCCGCTGGTCGACGCTCGACCTCTCGC	46380
	G L D L E A L A L P R R W S Q L D L S L	
46381	TGAGCATGGCGCGGCTCGGGACGGGCTGACGGGGGTGTGGGAGTACCGCACCGACCTGT	46440
	S M A R L G D G L T G V W E Y R T D L F	
46441	TCACCGAGGCCACGGTCGCGGAGCTGAGCGAGGCGTTCTCCACCTGCTGCGGGCGGCCG	46500
	T E A T V A E L S E A F V H L L R A A V	
46501	TCGAGGACCCGGGCGCGCCGCTGGAGACGCTGCCGCTCACCGGCGGCCGGGAGACCGGGC	46560
	E D P G A P V E T L P L T G G R E T G P	
46561	CGCGCCGCGGCCGTCGGCGGCCCGGCCCTCCCGCTGCACCGGCTCGTGGCCGCGG	46620
	R R G P S A A R P A L P L H R L V A A A	
46621	CGGCGCGCCGATCCCGCACGGACGGCGGTCTGCGACTCGCCCCGGACGGCACCGCCC	46680
	A R R D P A R T A V V A L A P D G T A H	
46681	ACCACATCAGCCACGGAGCCCTGCACCGCGCGCCACCACCCTCGCCGCGCGGCTCCGCC	46740

H I S H G A L H R A A T T L A A R L R R  
 46741 GGGAGGGCGCCGCGCCGAGCGGCCCGTCCGCTGCTCGTTCGAGCGGGGCCCTGGCTGC 46800  
 E G A G P E R P V A V L V E R G P W L P  
 46801 CCGTCGCCTACCTCGGCATCCTGCACGCCGGGGCCACCGTGTGCCCCCTGGACCCGGAGG 46860  
 V A Y L G I L H A G A T V L P L D P E D  
 46861 ACCCCCCGCACAGGCTCGCCCGGACGATCGCGAACTCGGGGGCGCGGCTGCTGCTCACCG 46920  
 P P H R L A R T I A N S G A R L L L T E  
 46921 AGACCGGGACCGCCTCGCGCGCGGCCGAGGCGGCCGCTCCGCGGTACGCGCGCTGACCG 46980  
 T G T A S R A A E A A G P G V R A L T V  
 46981 TGCCTGAGGGTGCCACCGCGCGGCGAGCGTTCTCGGCGGACGTCCACCCCGAGCAGTCCG 47040  
 R E G A T G G E R F S A D V H P E Q S A  
 47041 CGTACCTGCTGTACACCTCCGGGTGACGGGCGACCCCAAGGGCGTGTCTCGTCCCGCACC 47100  
 Y L L Y T S G S T G D P K G V L V P H R  
 47101 GGGCCATCGTCAACCGCCTCCTGTGGATGCAGGAGACCTACCGGTGCGCCCCGGGGGAGC 47160  
 A I V N R L L W M Q E T Y R L R P G E R  
 47161 GGGTCCTGCACAAGACGCCGGTGACGTTTCGACGTCTCGATGTGGGAGCTGCTGTGGCCGC 47220  
 V L H K T P V T F D V S M W E L L W P L  
 47221 TGACCGCCGGGGCGACCGTCTGTCATGGCCCGGCCCGGACCCACCGCGACCCCGCGCGAC 47280  
 T A G A T V V M A R P G T H R D P A R L  
 47281 TCGTCCGGCGGATCGCCCGCGAGGCCGTACACCGTGCACCTTCGTCCCCTCGATGCTCA 47340  
 V R R I A R E A V T T V H F V P S M L T  
 47341 CCCCCTTCCTACCGAGCTCGCCCGCGGCACGACGCGGTGCCCCGCTGCGGCGCGTGG 47400  
 P F L T E L A R G T T R L P A L R R V V  
 47401 TGTGCAGCGGGGAAGAGTGTCCCGCGGCCGCGTGAACCGCGCCCGGACTCCTCGACG 47460  
 C S G E E L P A A A V N R A A G L L D A  
 47461 CCCGGTGTACAACCTCTACGGCCCGACCGAAGCCCGCTCGACGTACCGCCTGGCCCT 47520  
 R L Y N L Y G P T E A A V D V T A W P C  
 47521 GCCGCCCCCGGAGCCGGGGCCGGTGCCGATCGGCCTGCCCATCGCCAACACCACCACCG 47580  
 R P P E P G P V P I G L P I A N T T T E  
 47581 AGGTCCTCGACGGCCGGCTGCGCCCGCTGCCCGCCCGGTGCCCGGCGAGCTGTACCTGG 47640  
 V L D G R L R P L P R P V P G E L Y L G  
 47641 GCGGCGCCTGCCTGGCCCATGGCTACCACCACGACCCGGCCCTGACCGCCGCGCGCTTCC 47700  
 G A C L A H G Y H H D P A L T A A R F L  
 47701 TTCCGGCCCCCGGCGGGCGCGCGCTACCGCACCGGGGACCTCGTCCGCAACGGGCGG 47760  
 P A P G G G R R Y R T G D L V R Q R A D  
 47761 ACGGGGCACTGGTGTTCGGGGACGCACGACGACAGGTGAAGATCGGCGGCATCCGGG 47820  
 G A L V F R G R T D D Q V K I G G I R V  
 47821 TCGAGCCCGGCGAGGTGGCGGAGGCGCTTCGGGCCCTGCCCGGCGTCCCGACGCGCGG 47880  
 E P G E V A E A L R A L P G V A D A A V  
 47881 TCGTCCCGCACGACGGGCGGCTGGCGGCGTACGCGGTGCGCGACCCGGTCCGCCCCGGCCC 47940  
 V P H D G R L A A Y A V A D P V G P A P  
 47941 CGGCGGCGGACGCCCTGCGGGACGCGCTGCGCAGGCGGCTGCCCGGCCACCTGGTGCCCG 48000  
 A A D A L R D A L R R R L P G H L V P A  
 48001 CCGCCCTCACCTGCTGGACCGGCTGCCCTCACCCCGCGGGCAAGCTCGACCGCCGGG 48060  
 A L T L L D R L P L T P A G K L D R R A  
 48061 CGCTGCCCCACCCGTGCGCCCCGCCCCGGACGGCGGACGGCCGCCACGACCGGGACCG 48120  
 L P H P S A P P P D G G R P P T T G T E  
 48121 AACGGCTCGTTCGCCCCGGGTGTGGGCCGAACGCTCGGACGGGAAGTCGTGCGCGTGGACC 48180  
 R L V A R V W A E R L G R E V V G V D R

48181 GGGACTTCTTCTCCCTGGGCGGCGACTCCGTCCGGGCCCTCGGCGTGACGGCGGCCCTGC 48240  
D F F S L G G D S V R A L G V T A A L R

48241 GCGCCCGCGGGCTCCCGGTGACGGTCACCGACCTCCTGCGCCTGCCACCGTGGCCGCCC 48300  
A A G L P V T V T D L L R L P T V A A L

48301 TCGCCCGCCACGCCGACGAGCGGGCGGATCGCCGACCGGCGGACAGGAGACGCCCCCG 48360  
A R H A D E R A D R R P A R Q E T P P G

48361 GGCCGTTTCGCCCTCTGCCCGGAAGCCGCCGGCGTGCCCGGCCTGGAGGACGCCTACCCGA 48420  
P F A L C P E A A G V P G L E D A Y P M

48421 TGTCGATGGCCCAGCGGGCCGTGCTCTTCCACCGTGACCACAACCCCGGTACGAGGTCT 48480  
S M A Q R A V L F H R D H N P G Y E V Y

48481 ACGTCACCAGCGTCGCGCTCTCCACGCCCCTGGACCGCACACGGCTCGCCGCGGCCGTGG 48540  
V T S V A V S T P L D R T R L A A A V D

48541 ACCGGTGCTGGACCGGCACGCCTATCTGCGGTCTCTTCGACCTCGTGTCCACCCGG 48600  
R L L D R H A Y L R S S F D L V S H P E

48601 AGCCACCCAGCTCGTCTGGACCCACCTGCCACCCCGCTCGAGGTGGTGGAGTCGTCCG 48660  
P T Q L V W T H L P T P L E V V E S S D

48661 ACCCCGCGGTTTCGACGCGTGCTGCACGCCGAACGCAAGCGCCCCCTCGACGTCGGCA 48720  
P A G F D A W L H A E R K R P L D V G T

48721 CCGGACCGCTGGCCCGGTTACCGCGCACGACGCGGGAGCCCGGATTCCGGCTGACCG 48780  
G P L A R F T A H D A G A A G F R L T V

48781 TCAGCAGCTTCGCCCTCGACGGCTGGTGCCTGGCCACCGTGCTACCGAACTGCTCCGCG 48840  
S S F A L D G W C V A T V L T E L L R D

48841 ACTACTGGTCCGCGCTGCGCGGCGGCCCCCTCAGCCTCCCGGCACCCGCGCCTCCTACC 48900  
Y W S A L R G A P L S L P A P A A S Y R

48901 GCGAGTTCGTCGCCCTCGAACGCGCCGCCAACACGATCCGGCGCACCGGGAGTTCTGGC 48960  
E F V A L E R A A Q H D P A H R E F W R

48961 GGACGGAGCTCGCCGGTGCCCGGCGCATCCGCTGCCCCGCGCCCGGTGCCACCGCCCG 49020  
T E L A G A R P H P L P R R P V P P P G

49021 GGCCGGACGGGATCCGCCAGCACCGTCACGTCGTCCCCGTGAGGACACCGTCGCCAAGG 49080  
P D G I R Q H R H V V P V E D T V A K G

49081 GCCTGTCCGCGCTCGCCGGCGAGCTGGGTGTGGGCTCAAACACGTTCTGCTCGGCGTCC 49140  
L S A L A G E L G V G L K H V L L G V H

49141 ACCTGCGGGTCGTCCGGGCCCTGTCCGGCGACCCCGACGTCATCACGGCCGTGGAGACCC 49200  
L R V V R A L S G D P D V I T A V E T H

49201 ACGGCCGCTCGAACGGCACGACGCGACCGCTCCTCGGGGTGTTCAACAACATCCTGC 49260  
G R L E R H D G D R V L G V F N N I L P

49261 CGTGCGGCAGCGGTGGACGGCGGGAGCTGGGCGACCTGGCCCGCGCCGCGCACGCGG 49320  
L R Q R V D G G S W A D L A R A A H A A

49321 CGGAGGCGCGGACGGGGAGTACCGCCGCTATCCGCTGGCCCAGGCACAGCGCGACCAG 49380  
E A R T G E Y R R Y P L A Q A Q R D H G

49381 GCGCGCCGGGCTCTTCGACACCCTCTTCGTGTTACCCACTTCCACCTCTACCGCGCGC 49440  
A A G L F D T L F V F T H F H L Y R A L

49441 TGGCCGACCTGGACGGCATGGCGGTCTCCGACCTGCGGGCCCCGACCAGACCTACGTAC 49500  
A D L D G M A V S D L R A P D Q T Y V P

49501 CGCTACCGCCCACTTCAACGTCGACGCCACGGACGGCGGCGCCTGCGGCTGCTGCTGG 49560  
L T A H F N V D A T D G G G L R L L L E

49561 AGTCGGACCCGCGGGAGTTCCTCCGACGAGCAGGTGCGGGAGTTCGCCGCGTACTACCGCC 49620  
S D P R E F P D E Q V A E F A A Y Y R R

49621 GCGCGCTGCGGGCCGCCCGGACGCCCCGACCGGCCGTACCGGGACACGCGGTTGACGG 49680  
A L R A A A D A P H R P Y R D T P L T D

49681	ACCGGCCGCGCGGTCCGGCGCCGACCGCGCGGAGCGCTCCGTCCACGCCCTGTTCCGCGG	49740
	R P A G P A P H R A E R S V H A L F A A	
49741	CCCCGGCCCCGAACCACCCGGACCGGATCGCGCTCGACGGCGAGGACGGGCGCGTCAAGC	49800
	P A R N H P D R I A L D G E D G P V S H	
49801	ACGGCGCCCTGGCCCCGGCGCGCCCGCCCGCTCGCCGGAACGCTGCGGGCGCGGGCGCCG	49860
	G A L A R R A A R L A G T L R A A G A G	
49861	GGCCGGACACCGTCGTGCGGATCTGGGCGCCGCGCCGCGCCGACGCGCTGTTGGCGCTGC	49920
	P D T V V G I W A P R R A D A V V A L L	
49921	TGGCCGCCCTCCACGCCGAGCCGCTACCTGCCCCCTGGACCCGGTCCACCCGCCCGGC	49980
	A A L H A G A A Y L P L D P V H P P R R	
49981	GGCAGCGGCAGGTGCTCACCGAGGCCGCGCCCGCTGCTCGTCTGCCCCGCGGCTCG	50040
	Q R Q V L T E A G A R L L V L P A G L D	
50041	ACACCCCGCTCCGGGCGCTGCGGCGCTGCGCGTCTGGCCCCGGACGACCTCGGCGCGCCCA	50100
	T P L R A C G L P V V A P D D L G A P I	
50101	TCGCCCCCGTGTCCGTCCACCCGGAGCAGCTGGCGGCGGTTCATGGCCACGTCCGCTCCA	50160
	A P V S V H P E Q L A A V M A T S G S T	
50161	CCGGGACGCCCAAGACGATCGGCGTCCCGCAGCGCGCCCTGGCCGGCTACCTCCGCTGGG	50220
	G T P K T I G V P Q R A L A G Y L R W A	
50221	CGATCGGCCACTACCGCCTCGACGAGGAGACCTCTCCCCGGTGCACCTCCTCGCTGGGCT	50280
	I G H Y R L D E E T V S P V H S S L G F	
50281	TCGACCTGACCGTCACCGCGCTGCTCGCACCGCTGGCCGCGCGGCGGCGAGGCGGGCTGA	50340
	D L T V T A L L A P L A A G G Q A R L T	
50341	CCGACTCCGGCGACCCGGGTGCCCCGCGCGGCACTGGCCGCGCGCCACACACCCCTGC	50400
	D S G D P G A L G A A L A A G H H T L L	
50401	TCAAGATCACCCCGGCCCATCTGGCCGCCCTCGCCACCAAGTTGGGCGCGCCGACCGCAC	50460
	K I T P A H L A A L A H Q L G A P T A L	
50461	TGCGCACCGTCTGTTGGCCGGGGCGAACCCCTGCACGCCGCGCCACGTCCGCGCCCTCCGCG	50520
	R T V V A G G E P L H A G H V R A L R A	
50521	CCTTCGCGCCCGCGCGCCGGCTCGTCAACGAGTACGGGCGGACCGAGACCACCGTCCGCT	50580
	F A P G A R L V N E Y G P T E T T V G C	
50581	GCTGTGCCACGACGTGCGACCGGACCCCGCGAGGCGCCCATCCCCGTGCGTACCCCGA	50640
	C A H D V A P D P G E A P I P V G T P I	
50641	TCGCGGGCCTCAGCGCGTTCGTGCTCGACGACGCGTGCCTCGCACCGCCCGCGTTCGGG	50700
	A G L S A C V V D D A L P A P P G V R G	
50701	GCGAGCTGTACATCGGCGGGACGGGCGTCACCCCGGCTACCTGGGCGGGCCCGCGGCCA	50760
	E L Y I G G T G V T R G Y L G R P A A T	
50761	CCGCCGCGCCTACGTGCGGACCCCTGCGCCCCCGCGCGCCGCTACCGCACCGGCG	50820
	A A A Y V P D P A A P G A R R Y R T G D	
50821	ACCTGGCACGCCGCTGCCGACGGCACCCCTGCTCCTGGCGGGGCGCGCCGACCGCCAGG	50880
	L A R R L P D G T L L L A G R A D R Q V	
50881	TGAAGATCCGCGGCCACCGGTGGAACCGGGGAGGTGAGCAGGTGCTCGGCGGCCACC	50940
	K I R G H R V E P G E V E Q V L G G H P	
50941	CCGGGGTGCGGGAGGCGGCGTTCGTGCGCCACCCGGCACCCGGCGGCGGCGCGGCTGG	51000
	G V R E A A V V A H P A P G G G R R L V	
51001	TCGCGTACTGGGTACCGGCCGAACCGGCCCGGCCACCGTCCGCGGACGCGCTCACCGCGC	51060
	A Y W V P A E P A R P P S A D A L T A L	
51061	TGCTCGCCGACCGGCTGCCGCGTACGCGGTCCCCCGGAACCTCGTCCGCTGCGCGCCC	51120
	L A D R L P P Y A V P A E L V R L P A L	
51121	TGCCCCACACCCCAACGGCAAGGTGACACACCCGGCTGCCCCGCGCGGACGGGACC	51180

	P	T	T	P	N	G	K	V	D	H	T	R	L	P	A	A	G	R	D	R	
51181	GGCGACTGGC	CGGAACTGCT	CGACCGGATCG	GAGGCACTGT	CCGACGCCG	AGCGGCCTCGG															51240
	R	L	A	E	L	L	D	R	I	E	A	L	S	D	A	E	A	A	S	A	
51241	CACTGCGCGA	CAGCCGGCCCCG	CAACCCGGGAGT	TGGCGATGACC	GAGCATGACG	ACCACCCG															51300
	L	R	D	S	R	P	A	P	G	S	G	D	D	R	A	*					
51301	CCGGCCCCGCC	GGGGCCCCGCCG	GTTCCGCTGG	CCCCGGCGGA	AGCCC GCCCGT	CCCCGCAC															51360
51361	GTGCCGGTG	CCCCGGGCATG	ACGACCCGCGT	CGGACGGCTG	CCGGCGGACC	GAGCGTCCCG															51420
51421	CCGACCCGCCG	ATTCTCTGGGG	ACCCCGCCG	GTTCCGGTGG	TGGCCGCCCGT	CCCCGCAC															51480
51481	CCGGAGGTG	CCGATGCGCGG	GACATGACGACC	CGCGTCGGAC	CGGCTGTTCG	GCGGACTGGAGC															51540 (orf21)
	M	R	G	H	D	D	R	V	G	R	L	S	A	D	W	S					
51541	GTCCCCGCCG	ACCCGCCTGCC	CGCCGGGGAC	CCCGCCGGTT	TCCGTCGGCC	CCCGCGGAGGC															51600
	V	P	P	T	R	L	P	A	G	D	P	A	G	S	V	G	P	G	G	G	
51601	CCGCCCCGT	TCCGCACGAGG	AGGTGACGATG	TCTCGGAGTAT	GACGACCGCCT	CGCGCGGCTG															51660
	P	P	V	P	H	E	E	V	T	M	S	E	Y	D	D	R	L	A	R	L	
51661	TCGGACAACC	CAGCGCGCCCT	GCTGGACCGCT	TGGCTCGCCG	AGGACCCCGC	CGGCGGTGCC															51720
	S	D	N	Q	R	A	L	L	D	R	W	L	A	E	D	P	A	G	G	A	
51721	GGCCCGCTT	CGCCCCGACG	GCGCCCGCCCG	CACCGAGGCC	GAGCGGATCCT	TGGCCGGG															51780
	G	P	L	R	P	D	G	R	P	P	R	T	E	A	E	R	I	L	A	G	
51781	GTCTGGGAGG	AGGTGCTGGAG	ACCGCGGGATC	GGCGCCGACG	ACGACTACTT	CGCGCTC															51840
	V	W	E	E	V	L	E	T	G	G	I	G	A	D	D	D	Y	F	A	L	
51841	GGCGGAGACT	TCCGTCCACG	CCATCGTCATC	GTGGCGAAGG	CCCCGGCAGG	CCGGACTCGCC															51900
	G	G	D	S	V	H	A	I	V	I	V	A	K	A	R	Q	A	G	L	A	
51901	CTGACCGCC	CATGACCTCTT	CGAGGCCAGG	ACCCTCGCGC	CGGTGGCGCG	GAGAGCCGCC															51960
	L	T	A	H	D	L	F	E	A	R	T	L	A	A	V	A	R	R	A	A	
51961	CCGGCCG	CCCCGCGAG	CCCGTCCCCG	ACGCGGGCGG	CGCGCGGTCC	CGTACCCGCTG															52020
	P	A	G	P	A	E	P	V	P	D	A	G	G	G	A	V	R	Y	P	L	
52021	ACCCCTATG	CAGCAGGGCAT	GCTCTACCACT	CGGCCGCGGC	CAGCACGCC	CGCGCCTAC															52080
	T	P	M	Q	Q	G	M	L	Y	H	S	A	G	G	S	T	P	G	A	Y	
52081	GTGGTGCA	GGTGTGCTGCC	GGGTGACGGG	GGACCTCGAC	GTGGCCGCTT	CCGCACCGCC															52140
	V	V	Q	V	C	C	R	L	T	G	D	L	D	V	A	A	F	R	T	A	
52141	TGGCAGG	CCGTGCTGTCC	GCCAACCCGG	CGCTGGCCGT	CTCTCTTCC	ACTGGTCCG	ACGGC														52200
	W	Q	A	V	L	S															





	L A R S D V A E P V N I G S E E R V D I	
54181	TCGCGTCGCTCGTCGAGCGGATCGCCGGGGTTCGCCGGGAAGAAGGTGCGCTGCGCCTTCG A S L V E R I A G V A G K K V R C A F A	54240
54241	CCCCGACCGCCCGGTTCGGGCCCCCGGGCGCGTCTCGGACAACACCCGCTGCCGCGAAC P D R P V G P R G R V S D N T R C R E L	54300
54301	TGCTCGGTGGGACCGGAGACGTCCCTCGCGGCCGGCCTGGAGCGCACCTACCCGTGGA L G W A P E T S L A A G L E R T Y P W I	54360
54361	TCGAGCGCCAGGTCCTCGCCGAGGCCGGGAGGGCCGATGCCTGAGCACCGCACACCGGTG E R Q V L A E A G R A D A *	54420 (orf19)
54421	AAGGACCTCGGCCGGTCTGCTCGGGCACGCCGCGCTTCCGGGGCCGCGAGCTGCAG K D L G R L L L G H A A R F R G R E L Q	54480
54481	GACGTCGCCACCCGGGCGCTGCGGGCCTCCGGCGGGGAGAACGCCTGGGTGGTGTCCGTC D V A T R A L R A S G G E N A W V V S V	54540
54541	GTCAACACCAGTCTCCGCGCCCGCCAGGCCGTGGACCACGCGCTGCGGCTCGCCCCCGC V N T S L R A R Q A V D H A L R L A P R	54600
54601	CGCGGGCTCTCCCGGTGCGTACCCGTTCTCCGCCGCCACCACACGGCCACCCGCCC R G L S R L R Y P F S A A H H T A T P P	54660
54661	CGGACCTGTGCTGCTGTGCCCCAGCCGCGAACGCGTCGGCAACGTCGAACGCTTCCTC R T L S L L C P T R E R V G N V E R F L	54720
54721	GACAGCGTCGCCCCGACCGCCGCGCGCCGGCGGATAGAGGCCCTCTTCTACGTCGAC D S V A R T A A A P G R I E A L F Y V D	54780
54781	GACGACGACCCCCAACTCCCTGCCTACCACGAGCTGTTTCGAGCACGCCCGGTGGCGCTAC D D D P Q L P A Y H E L F E H A R W R Y	54840
54841	GGACGGATCGGCCGGTGCGCCCTGCACGTTCGGCGCCCCCGTCGGCGTACCCACGCCTGG G R I G R C A L H V G A P V G V P H A W	54900
54901	AACCACCTGGCCCGAACGCGGCCGGCGACGTGCTGATGATGGCCAACGACGACCAGCTC N H L A R N A A G D V L M M A N D D Q L	54960
54961	TACATCGACTACGGCTGGGACACCGCCCTCGACGCCCGCTCACCGAACCTGAGCGCCCTG Y I D Y G W D T A L D A R V T E L S A L	55020
55021	CACCCCGACGGCGTCTGTGCCTGTACTTCGACGACGGCCAGTACCCCGAGGGCGGCTGC H P D G V L C L Y F D D G Q Y P E G G C	55080
55081	GACTTCCCGATGGTGACACGGCCCTGGTACGGCACCCCTCGGCTACTTCACCCCGACGATC D F P M V T R P W Y G T L G Y F T P T I	55140
55141	TTCCAGCAGTGGGAGGTGAGAAAGTGGGTCTTCGACATCGCCGACCGGCTGCACCGGCTC F Q Q W E V E K W V F D I A D R L H R L	55200
55201	TACCCCGTCCCGGCGTCTCTGTCGAACACCGGCACTACCAGGACTACAAGGCACCCCTTC Y P V P G V L V E H R H Y Q D Y K A P F	55260
55261	GACGCCACCTACCAGCGGCACCGGATGACACGGGAGAAGTCTTCGCCGACCACGCCCTG D A T Y Q R H R M T R E K S F A D H A L	55320
55321	TTCCTGCGCACCGAGCCGGACCGCGAGGCGGAGACGGACAGGCTGCGGGCCGTCATCGCC F L R T E P D R E A E T D R L R A V I A	55380
55381	CGGGCAGGGAACACCCCGACGCCGACCACGCCGACCATGCCGTTACGACGCGGAGACC R A G N T P D A D H A D H A V H D A E T	55440
55441	TTCTGGTTACCGGCCTCTGCGCGAGTCCCACGCCAAGCTGCTCGCGGAACCTCGACGAC F W F T G L L R E S H A K L L A E L D D	55500
55501	GCGCCGGGCCCGGCCGCGAGCCGTGCTCTTCGCCGACGGCTCCTGGACCGCGTTCGCC A P G P A A G A V L F A D G S W T G V A	55560
55561	TACCGCACCCACCGCTGGCCACCGCCCTGCTCGCCTCGATCCCCGAGGCCACCCCTCGAC Y R T H P L A T A L L A S I P E A T L D	55620

55621	TCCGGCCGCGCCGACCTCCTCGTCGTCCCGCCGCGCGTCCCACCACCACCCCGACGGC	55680
	S G R A D L L V V P P G A S H H H P D G	
55681	ACCGTCGACTCCGCGTTCGCTCCGACGCCGCGCTCCGCGTCTGTTCGGACTGCGCGTG	55740
	T V D S A F G S D A G L R V L F G L R V	
55741	CCGGACGCCGCGCAACTCCGCGTCGGCGACGGCCCGGTGCCCTGGGGCAATGGGCAATGC	55800
	P D A A Q L R V G D G P V P W G N G Q C	
55801	CTGATCCACGACACCGCCGACCGAGCACCTGCGCAACGACGGCACCGAATCTCTGGCC	55860
	L I H D T A A P S T L R N D G T E S L A	
55861	GCCCTCACCTTCGTGGTCCCGCGCCCGGCACCGGGGAGTGAGGCCGTGTGCGGCATCG	55920
	A L T F V V P R P A P G E *	
	M R P V C G I V (orf18)	
55921	TGGCGATCCGCTCCGCCGACGGCGGACTCGACGGCGGTGAACTCACCGCGCCGATGGCCG	55980
	A I R S A D G G L D G G E L T A P M A D	
55981	ACCTGCGCCCGCGCGGCCCGACGGCGAAGGCACCTGGGTCTCGCCACCGGCCGGGCCG	56040
	L R P R G P D G E G T W V S P T G R A A	
56041	CCCTCGGCCACACCCGGCTCGCCGTGATCGCCCCGACCGCGACGCCAGCCGGTCGCCG	56100
	L G H T R L A V I A P D A G R Q P V A G	
56101	GCCCGGACGGCACCGTCCGCTCGTCGTCAACGCGAGTTCTACGGCTACCGGGAGATCC	56160
	P D G T V R L V V N G E F Y G Y R E I R	
56161	GCGCGGAAGTGC GCGCCGCGGCTGCGCGTTCCGCACCGGCAGCGACAGCGAGATCGCCC	56220
	A E L R A A G C R F R T G S D S E I A L	
56221	TCCACCTGTACCTGCGGGACGGCCGGCGGGCACTGGAGCGGCTGCGCGGCGAGTTTCGCT	56280
	H L Y L R D G R R A L E R L R G E F A F	
56281	TCGTCTCTGGGACGAACGCCGCGCCACCCTCTTCGCGCCCGCGACCGGTTTCGGCGTCA	56340
	V L W D E R R A T L F A A R D R F G V K	
56341	AACCCCTCTACTACACCGAGCGCGACGGCGGGCTCTACGTGCGCTCGACGGTCAGGGCCC	56400
	P L Y Y T E R D G R L Y V A S T V R A L	
56401	TGCTCTCTGCGGGCGCCCCGCGCGCTGGGACACCGCGCCTTCGCGCGGCACCTGCAGC	56460
	L S C G A P A R W D T A A F A A H L Q L	
56461	TCGGCCTGCCCCCGACCGCACCTCTTCGCGCGCATCCGGCAGCTCCCGCCCGGCTGCC	56520
	G L P P D R T L F A G I R Q L P P G C H	
56521	ACCTCATCGCCGACGCCACCGCACCCGCGTACCCCTACTGGGACCTCGACTACCCGC	56580
	L I A D A H G T R V T P Y W D L D Y P P	
56581	CCGCCGCGAAGTCCGCCGCGGGGAAGCTGGACGACCACCTGGACGCGGTACGCGAAC	56640
	A G E L A A R G S L D D H L D A V R E R	
56641	GGACCGACGAGGCCGTACGTTGCGTACCGTCGCCGACGTGCCCTCGCCTGCCACCTCA	56700
	T D E A V R L R T V A D V P L A C H L S	
56701	GCGGCGGCTGGACTCCTCCGCCGTGCGCGCTCCGCCGCGCCACACCCGGCTCACCG	56760
	G G L D S S A V A A S A A R H T R L T A	
56761	CCTTCACCGTCCGCTTCGACGACCCCGCCTTCGACGAGAGCGCCGTGCGCCGGCGCACCG	56820
	F T V R F D D P A F D E S A V A R R T A	
56821	CCGCCACCTGGCCATCGACCACCGCGAAGTCGCTCGGAACGCGCCCACTTCGCGGACC	56880
	A H L A I D H R E V A S E R A H F A D H	
56881	ACCTGCGGGACGTGTCGCGCCCGCGAGATGGTGAGGAGAACTCGCACGGCATCGCCC	56940
	L R D V V R A G E M V Q E N S H G I A R	
56941	GGTACCTGCACAGCGCGCATCAAGAAGGCGGGATTACCGCCGTCTCGCCGGGGAGG	57000
	Y L H S A H I K K A G F T A V L A G E G	
57001	GCGGGGACGAAGTGTTCCTCGGCTACCCCACTTCGCAAGGACCTGACGCTCAGCCTGT	57060
	G D E L F L G Y P Q F R K D L T L S L S	

57061	CCGCCGACGCCCCGCGACAAGGCCGACCGCGGTACGCCCGGCTGGTCGCGGCCGGGCTCC	57120
	A D A R D K A D R G Y A R L V A A G L L	
57121	TGCCGCGGTACCTGCGCACCTCTCTGGCACCTCGGCTTCTTGGCTTCTGGATCGTCG	57180
	P P Y L R T L L G T L G F L P S W I V D	
57181	ACCGCCACCTGGCCGTCACCCAGCCCGTCGCGCCCTGCTCCGCCCCGACTTCGCCGCCG	57240
	R H L A V T Q P V A A L L R P D F A A E	
57241	AACTGGCCCCGCGCCGACGCCCGCGCCCCCTGCTCGCGCCGGCGCCGGCCTGCTCGCCG	57300
	L A R A D A A A P L L A A G A G L L A G	
57301	GGCGCGCCCCGGCGCACCCAGGCCACCTACCTCTTCGCAAGACCTGGCTGCCCCGGCTACC	57360
	R A P A H Q A T Y L F A K T W L P G Y L	
57361	TGCTCGCGCCCGAACGCCTCGACGCGGCCAGGCCGTCGAGGTGCGGCTGCCCTCTTCG	57420
	L A A E R L D A A Q A V E V R L P L F D	
57421	ACCACCACCTCTTCGACCTCGTCCGGCACACCCCGCGCCCTGGTACGACAAGGACGGCA	57480
	H H L F D L V R H T P P A W Y D K D G T	
57481	CCGGCAAGTACCCGCTGCGGGCCGCCATGCCCCACCGGCTGCCGCGGAGGTGACCGAGG	57540
	G K Y P L R A A M R H R L P R E V T E G	
57541	GCCGCAACAGGGCTTCTTCGCACCTCCGATGGCCGACGACACCCCTCCTCGACGCC	57600
	R K Q G F L A P P M A D D D T L L D A L	
57601	TGCGCGAACGCCTCGCCGGACCGGGCGGGCGACGACCCCTTCTTCGACCCGACGCCG	57660
	R E R L A G P G A G D D P F F D P H A V	
57661	TCCGCGCCTGCTGGACCGCTGGCCGCCGACCCCCGGGCGAGCGGTCCGGCGGCGAGA	57720
	R A L L D R L A A A P P G Q R S G G E K	
57721	AACTCTCCAACCTCGTCGCGAGCACCGCCGAACCTGGCCGACGAGTTCGGCCTCACCACCG	57780
	L L Q L V A S T A E L A D E F G L T T A	
57781	CCCCAGCGGGCAGAAAGGCGGCAACGGTGGCTGACCTCGATCCCGGCACGCTCTCCGAG	57840
	P S G Q K G G N G G *	
57841	GCCGAGCTGACCGCCCGGATCGCGCCCTGTCCCCGAACGCCGGGCGGCTTCGAGAAG	57900
57901	ATGCTGCACGGCGCCGCGCACCCCCGCCCCGGCATCCCGCGCCGCGGCCACCGCGGCA	57960
	M L H G A A H P R P G I P R R G A T A A	(orf17)
57961	CCGGCCTCTACGCCAGGAACGCCTGTGGCTGCTCACCGGGCTGCTGCCACCGCCTAC	58020
	P A S Y G Q E R L W L L T G L L P T A Y	
58021	AACTACGCCACCGCCCTGCGGCTGCGCGGCGACCTGTCCGTCCCCGCGCTGCGGGCGCC	58080
	N Y A T A L R L R G D L S V P A L R G A	
58081	CTGCGCGCATCGTCCGCCGCCACGAGGTGCTGCGCACACCTTCCGGCTGGACGGCGAC	58140
	L R G I V R R H E V L R T T F R L D G D	
58141	GACCTCATCCAGGTGCTCCACCCACGGCGGACGTCCCCGTGCGCCTGGCCGACCTCACC	58200
	D L I Q V V H P T A D V P V R L A D L T	
58201	GGACGCTCCGCCGACACCGGGCGGCTGATGCGCGAGGAGGCCCGCCGCCCTTCGACCTG	58260
	G R S A D T G R L M R E E A R R P F D L	
58261	GAGCACGGGCGCTGCTGCGGCTGACCTCTTCCGGCTCGGCCCCCGGACACCTCGCC	58320
	E H G P L L R L T L F R L G P R D H L A	
58321	CTGCTGGCGTCCACCACGCGTCCACCGACGGTGGTCCAACGGCGTCTCGTGACCGAA	58380
	L L A V H H A V T D G W S N G V L V T E	
58381	CTCGCCACCGGCTACCGGGAACCTGCGCGCCGACGCCCCGACGGCGGCCCGCCCCCGG	58440
	L A T G Y R E L R A G R P D R R P A P P	
58441	GTCCAGTACGGCGACTACGCGCACTGGCAGCGCGAGCGGCTGACCGGGCCGAACCTGCGG	58500
	V Q Y G D Y A H W Q R E R L T G P E L R	
58501	GCCCTGGAGGACTACTGGCGCACCGCCGTACGCGACCTGCCAGGACGGACCTGCCACCC	58560
	A L E D Y W R T A V R D L P R T D L P T	

58561	GACCGCCCCCGCCCCGCGCCCGGGCGCGGCGAGGGCGCCAACCACGCCCTGCTGCTCTCG	58620
	D R P R P A A R R G E G A N H A L L L S	
58621	CCGGAGCTGACCGGCCGGCTCGCCGACCTGCGCCGCGAGGGCGGGTCTGCTGTTTCATG	58680
	P E L T G R L A D L R R R E G G S L F M	
58681	CTCGTGTCTCCGCGCTCCTGGTCTGCTCCTGCGTGACCGGCGCGGGACCGGCTCGCC	58740
	L V L S A L L V V L R G T G G R D R L A	
58741	GTCGGCACCTCTGTCGCCGGCCGACCCGCCCCGAACCTCGAGCCGCTCATCGGCTACTTC	58800
	V G T L V A G R T R P E L E P L I G Y F	
58801	GTCAACGTCTGCTGCTGCCCTTCGAGACCGGCGCGCCGACCTCCTTCGCCGAGCTGTGG	58860
	V N V L L L P F E T G G R T S F A E L W	
58861	CGGCGGGTCCGCGGCCGGCTGGTGGAGGCGTACGCCCACCAGGAACCTGCCGCTGGAGAAG	58920
	R R V R G R L V E A Y A H Q E L P L E K	
58921	GCCCTGGAGCTGCTGCGCGCCGACGGCACCGCCCCCGCGACCCGCGGTCGGCGTGGTC	58980
	A L E L L R A D G T A P A D P P V G V V	
58981	TGCGTCGCCCAGCAGCCCCGCGGATCACCTGCGCGGACTCGACGCGAGCGTCGAG	59040
	C V A Q Q P A P A I T L P G L D A S V E	
59041	GACGTCGACCTGGGCACCGCCAGTTCGACCTCGTCGTCGAGGTGCGGAACGGCCGGA	59100
	D V D L G T A Q F D L V V E V R E R P E	
59101	GGCGTGCAGATCGCCTTCCAGTACGACCGGGACCTGTTTCGACGCGGCCACGGTCCGGCTC	59160
	G V Q I A F Q Y D R D L F D A A T V R L	
59161	CTCGCCGACCACGTGCACGCGCTCCTCGACCGCGCGCGCCGACCCACCTGCCCTGT	59220
	L A D H V H A V L D Q A A A D P T L P C	
59221	GCCGAGCTGCCCCGCCCCGCGCCCCCGCGCCCCGCGCCGACGGCCGGCGCCACGACG	59280
	A E L P A P P A P A A P A R T A G A T T	
59281	CTGCACGCCCTGTTTCGAGTCCCGCGCCGCGAAGAGCCCCGACGCGGTGCGCCCTCGTCGAC	59340
	L H A L F E S R A A K S P D A V A L V D	
59341	GGCGGCCACCGGTCACCTACCGGACCTCAACACCGCGCCAACCGGCTCGCCCGCCAC	59400
	G G H R V T Y R T L N T R A N R L A R H	
59401	CTGCGCGCGGTGCGCGTGCTACCGAGGACCGGTGGCGCTGCGCCTGCCCGCGGCACC	59460
	L R A V G V R T E D R V A L R L P R G T	
59461	GACGCGGTGACCGCCACCCTCGCCGCCCTCAAGGCCGGCGCGCTACGTACCCCTCGAC	59520
	D A V T A T L A A L K A G A A Y V P L D	
59521	CCCGCCCTCCCCGAGGAACGGCTGACCCGCGTCTCGCCGACGCCCGCCCGCGTGGTC	59580
	P A L P E E R L T R V L A D A R P A V V	
59581	CTACCCCCCGGTATCTGCACGACCGGTCCGCCGAGATCACCGCCACGCCGGCCATGAC	59640
	L T P A Y L H D R S A E I T A H A G H D	
59641	CTCAACCTCCCCGTCCACCCGACAACCTCGCCTACCTCCTCCACACCTCCGGATCCACC	59700
	L N L P V H P D N L A Y L L H T S G S T	
59701	GGCACCCCCaAGGGCGTCTCGGCAcCCACCGGGCGCGGTCAACCGGTGACTGGATG	59760
	G T P K G V L G T H R G A V N R V D W M	
59761	AGCACCGGTACCCGTTCCGGACCGGCGACGTGGCCGTGCGCCGACCGCGCCCGGCTTC	59820
	S T A Y P F R T G D V A V A R T A P G F	
59821	GTCGACGCGGTCTGGGAACCTCTCGGCCCCCTGGCCGCGGCGTCCCCCTCGTCCTCCTG	59880
	V D A V W E L F G P L A A G V P L V L L	
59881	CCGACCGACGAGGCGCGGACCCGGCCCTGCTGACGGCGGCGTGAACGGCACCGGGTG	59940
	P T D E A R D P A L L T A A L E R H R V	
59941	AGCCGGATGGTGACGGTCCCGTCTGCTGACCATGCTCCTGGACGAGTCCGCCCCGCGG	60000
	S R M V T V P S L L T M L L D E S A R A	
60001	ACGGACCTCGGCACCCGCTGGCCTGCCTCCGCACCTGGATCACCAGCGGCGAGCCCCTG	60060
	T D L G T R L A C L R T W I T S G E P L	

00442660

60061	CCGCCCCGCTCGCCCGGCGTTCCACGACCGCTGCCCGGCCGACCCTGCTGAACCTG	60120
	P P A L A R R F H D R L P G R T L L N L	
60121	TACGGCTCCTCCGAGACCGCCCGACGCCACCGCGGCCCGCATCGACCCGGCGCCCGG	60180
	Y G S S E T A A D A T A A R I D P A P G	
60181	ACTGCGCTCCCGGAGCGGTCCCGATCGGCACGCCCATCACCGGCGTCAGCGCCCTCGTC	60240
	T A L P E R S P I G T P I T G V S A L V	
60241	CGCGGCCCGGACCTGCGCCCGTGCCTCGCGCTGATGCCCGGCGAGCTGTACGCCGGGGC	60300
	R G P D L R P L P A L M P G E L Y A G G	
60301	GCGTGCGTGGCCCGGGCTACCACGCCCGTCCGGCCGAGACCGCCGCGGCGTTCCCGCCG	60360
	A C V A R G Y H A R P A E T A A A F P P	
60361	GATCCCGACGGCGGGCCCGCGCCCGGATGTTCCGTACCGGTGACAGGGCCCGGCTGCGG	60420
	D P D G G P G A R M F R T G D R A R L R	
60421	GCCGACGGCCGGCTGGAACCTCTGGGGCGCGTGGACCGGCAGGTGCAGATCCGCGGCCAG	60480
	A D G R L E L L G R V D R Q V Q I R G Q	
60481	CGCGCCGAGCCCGGCGAGGTGGAACACGCCCTGCTGGCCACCGCGCGGTACGGGCCGCC	60540
	R A E P G E V E H A L L A H P A V R A A	
60541	GCCGTACGGCGAACCCCGACGCCACCGGCCTGTGGGCGTACGTGCGGCTCGCTCCCGGC	60600
	A V T A N P D A T G L W A Y V R L A P G	
60601	CCGTTGCGCGCGGCTCCCCCAGACCGAGCTGACCGCCTTCTGCGCCGACGCTCCCT	60660
	P F A A G S P Q T E L T A F L R R T L P	
60661	GCCACCTCGTGCACCGCGGTACCGTCTGGACGAGCTGCCGGTGACCGCGCACGGC	60720
	A H L V P T A V T V L D E L P V T A H G	
60721	AAGACCGACCACGCGCGGTGCCCCCCCCGACCCCGGGCGGGCGCCCGCCCCGACC	60780
	K T D H A R L P A P D P R A G R P A P T	
60781	GCCCCCGCACCCCCACCGAGCGTACGGTCCCGACGTCTTCGCCGGGGTGCTCGGCCTG	60840
	A P R T P T E R T V A D V F A G V L G L	
60841	GAGGGGCGGGTCCGCGCGCACGACGACTTCTCTCTCGGCGGGCACTCCCTCCTCGCC	60900
	E G P V G A H D D F F L L G G H S L L A	
60901	GCCCGCAGTCGCGGCGGAATCCGCGCCCGCGCGGCGTCCGGATCGGGCTGAGCGACGT	60960
	A R S R G G T P R P P R R P D R A E R R	
60961	CTTCGCGGCCCCACCGTCCGCGCAGCGTCCCGCCCGGACCGACCGCCCGGCGCCCGG	61020
	L R G P H R R R S V A A R T D A A R P G	
61021	ACCGGCCCCGAGCACACCCCGTTCGTACCGACCCCGGCGCCCGGCACGAGCCGTTC	61080
	T G P E H T P F V T D P G A R H E P F P	
61081	CTCACCGAGTCCAGCGGGCTACTACGTGGGACGCGAGGGCGGGTTCGCCCTCGGCGGC	61140
	L T D V Q R A Y Y V G R E G G F A L G G	
61141	GTCTCCACCACGCTACCTGGAGATCGAGGCCCGCGGATCGACGTCGACGGTTTACC	61200
	V S T H A Y L E I E A P R I D V A R F T	
61201	GGCGCGTGCGCGGGGTGATCGCCCGCACCCCATGCTGCGCGCCGTGATCCGTCCCGAC	61260
	G A L R G V I A R H P M L R A V I R P D	
61261	GGGCTCCAGCAGGTGCTCACCGACGTCCCCCGTACGACGTGGCCGTGCACGACCTGCGC	61320
	G L Q Q V L T D V P P Y D V A V H D L R	
61321	GACCTGGACGAGCCCGCGCGCAGCGCCGACGCGCGCTGCGCGAGGAGATGTCCAC	61380
	D L D E P A R Q R R R A A L R E E M S H	
61381	CAGGTGGTGCCCGCGACCTCTGGCCCCGTTCGACGTCCGCGTCTCCCTCGGCCCCACG	61440
	Q V V P A D L W P L F D V R V S L G P T	
61441	GACGCCCTCGTCCACGTGGGGGTGGACGCGCTGATCTGCGACGCCACAGCTTCGGCCTC	61500
	D A L V H V G V D A L I C D A H S F G L	
61501	GTCCTGGCCGAACCTCGCGGCCCGTTACGCCGACCCCGCACGCGCTTCCCGCCCCCTGACG	61560

	V	L	A	E	L	A	A	R	Y	A	D	P	A	R	R	F	P	P	L	T																											
61561	G	C	G	G	A	C	T	T	C	C	G	G	A	C	G	T	C	C	G	G	A	A	C	C	G	C	G	A	G	T	A	C	G	C	G	61620											
	A	D	F	R	D	H	V	L	H	Q	E	A	L	R	G	T	A	E	Y	A																											
61621	G	C	G	G	C	G	A	G	C	G	T	A	C	T	T	C	C	G	A	A	C	C	G	C	C	G	C	C	G	A	A	C	T	G	C	C	61680										
	A	A	E	R	Y	W	R	E	R	L	P	E	L	P	P	G	P	E	L	P																											
61681	C	T	G	G	C	C	T	C	G	C	C	C	G	A	G	A	C	C	C	T	C	G	G	C	C	C	G	C	T	C	C	G	C	C	G	T	G	61740									
	L	A	V	A	P	E	T	L	G	T	P	R	F	T	R	R	S	G	R	L																											
61741	G	A	C	G	C	G	C	T	C	T	G	G	A	C	G	G	C	C	C	G	C	C	G	C	C	G	C	C	G	G	G	C	T	C	A	G	C	C	C	T	C	61800					
	D	A	A	S	W	T	A	V	K	D	R	A	R	R	A	G	L	S	P	S																											
61801	G	G	C	G	T	A	C	T	G	C	T	G	G	C	G	C	G	T	T	C	G	C	C	G	A	G	C	C	G	C	C	G	C	C	G	C	T	A	C	61860							
	G	V	L	L	A	A	F	A	E	V	I	T	A	W	S	G	R	P	R	Y																											
61861	T	C	G	C	T	G	A	T	G	C	T	G	A	C	G	C	C	C	G	C	C	T	C	C	A	C	C	G	A	C	T	C	G	G	C	G	A	T	C	G	T	61920					
	S	L	M	L	T	V	F	D	R	P	P	L	H	P	D	L	G	R	I	V																											
61921	G	G	C	A	C	T	T	C	A	C	T	C	G	T	C	A	G	C	T	G	G	A	G	T	C	G	A	C	C	G	C	G	A	C	T	T	C	A	C	61980							
	G	D	F	T	S	L	S	L	L	E	V	D	H	S	R	P	G	D	F	T																											
61981	G	A	C	A	G	G	C	C	C	G	C	C	C	T	C	A	G	C	G	C	C	T	G	T	G	G	C	A	G	A	C	T	C	G	A	C	C	T	G	G	C	T	62040				
	D	R	A	R	A	L	Q	R	R	L	W	Q	D	L	D	H	L	A	V	G																											
62041	G	G	C	G	T	G	A	C	G	T	G	A	C	G	G	A	A	C	G	G	G	C	G	T	G	C	C	A	C	G	C	C	C	G	A	C	C	C	G	T	G	T	C	A	C	A	62100
	G	V	T	V	T	R	E	R	A	L	R	H	D	A	R	P	G	L	L	T																											
62101	C	C	C	G	T	C	G	T	C	T	T	C	A	C	T	C	C	G	A	C	T	G	C	T	G	T	C	G	G	C	G	A	C	C	G	C	C	G	A	C	C	G	A	C	C		

63001	GGCGACGCGGTCTCTCGCCGTCTCTCCCCGAGCTTCGACCTCGCCGTCTACGACCTGTTTC	63060
	G D A V L A V S S P S F D L A V Y D L F	
63061	GGCGTGCTGGCCGCCGGCGGCACCGTGGTTCGTCCTCCCGCCACGACCGCCGGCGGACCCC	63120
	G V L A A G G T V V V P A H D R R R D P	
63121	GGACACTGGGCCGAGCTGATCCGGCGCGAGCGGTCACCCTGTGGAACCTCCGTCCCCGCG	63180
	G H W A E L I R R E R V T L W N S V P A	
63181	CTGGGCACCCTGCTCACCGAGTACGCCGAGGCCCTCGCCCCGACGCCCTGCGCACCCCTG	63240
	L G T L L T E Y A E A L A P D A L R T L	
63241	CGGGCGGTGCTCTCTCAGCGGCGACTGGATCCcctcgactgccccgaccGGATCCGCGCC	63300
	R A V L L S G D W I P L G L P D R I R A	
63301	CTGTCCGCCCCCGGCGCCACCGTGATGAGCCTCGGCGGCGGACCGAAGCCTCCATCTGG	63360
	L S A P G A T V M S L G G A T E A S I W	
63361	TCGGTCTGGTACGAGATCGGGAAGGTGCACGAGGCGTGGAGCAGCATCCCCTACGGCACC	63420
	S V W Y E I G K V H E A W S S I P Y G T	
63421	CCCATGGCCAACCAGCGGCTGGAGGTCTCGACGAGCAGCTGCGGCCCCGGCCCCGACTGG	63480
	P M A N Q R L E V L D E Q L R P R P D W	
63481	GTGCCCCGGGAGCTGTACATCGGCGGCACCGGCGTCGCCAAGGGTACTGGCGCGACCCG	63540
	V P G E L Y I G G T G V A K G Y W R D P	
63541	GAACAGACCTCCCTGCGCTTCCCCGTCCACCCGGGCGAGCGGGCAACGCCTGTACCGCACC	63600
	E Q T S L R F P V H P G S G Q R L Y R T	
63601	GGGGACTTCGCCCCCACCTCCCCGACGGCACGCTGGAATTCTGGGCCGGCAGGACGAC	63660
	G D F A R H L P D G T L E F L G R Q D D	
63661	CAGGTGAAGATCGGCGGATTCCGGGTGCAACTGGGCGAGGTGAGGCGGCCCTCGGCCGA	63720
	Q V K I G G F R V E L G E V E A A L G R	
63721	CTGCCCCGACGTCGCCCGCCGGCGCGGTGATCGCCACCGGTGACCCGCGGGGCGACCGCCGC	63780
	L P D V A A G A V I A T G D P R G D R R	
63781	CTCGTCGGCTTCGCCGTACCGGCCCGGGAGGGCGGCTTCGACGCGGCCGGGCTCCGACGG	63840
	L V G F A V P A R E G G F D A A G L R R	
63841	CAACTCGCCCCGGCGGCTGCCCGCTACATGGTCCCCACGACCTGTGCCCCCTGGACCGG	63900
	Q L A R R L P A Y M V P T T L L P L D R	
63901	CTGCCGCTGACCGCCAACGGCAAGGTGACCGGGCCGCACTCCAACGCCTCGTCCCCGGC	63960
	L P L T A N G K V D R A A L Q R L V P G	
63961	CGCGCACCGGCCCCGGCGGAACCCGCCACCGCCCCACCTGCCCCGTTCGCGCGCGTCCCC	64020
	R A P A P A E P A T A P P A R S R A V P	
64021	GTGCCCCGCTGGCTCGCCGACCTGTGGTGCAACTCCTCGACGTGCCGGAGGCCGACCCC	64080
	V P G W L A D L W C E L L D V P E A D P	
64081	GACGCGAACTTCTTCGCCCTCGGCGGCACCTCCCGGGTCGCGATCACCTGGTCACCCGG	64140
	D A N F F A L G G T S R V A I T L V T R	
64141	ATCGAGGCCCGACTCGCCGTCCGGGTGCCCTCGCCCGCTCTTCGACGCCCGCACCCCTG	64200
	I E A R L A V R V P L A R L F D A R T L	
64201	GGCGGCCTCGCCGAGACGATCGCCGAACCTGTGCGCCGCCCGGAGGAGGAGCCGGCACCC	64260
	G G L A E T I A E L S A A A E E E P A P	
64261	GCCGAGCCCGTGTACGCCCCGACCCCGCCACCCGCCACGAGCCGTTCGCGCTCACCGAC	64320
	A E P V Y A P D P A T R H E P F P L T D	
64321	ATCCAGCGCGCTACTGGCTCGGCCGGCACCGTCCCTCTCCCTTGGCGGCGTCGCCACG	64380
	I Q R A Y W L G R H R S L S L G G V A T	
64381	CACACCTACCTCGAACTCGACGTGAGGACCTCGACCCCGCGGCTCCAGACGGCCCTC	64440
	H T Y L E L D V E D L D P G R L Q T A L	
64441	CGCCGGCTGATCGACCGCCACGACGCCCTCCGGCTCGTGGTCTCCCCGACGGCCGGCAA	64500
	R R L I D R H D A L R L V V L P D G R Q	

64501 CAGATCCTCGGCGACGTACCGCCGTACCTCCTCGCCACACCGACCTGCGGGCAGGGCG 64560  
Q I L G D V P P Y L L A H T D L R G R A

64561 GACGCCGAGGCCGAAC TGGCCCGCGTCCGCGAGCACATGTTCGCACGAGGTGCGCGACGCC 64620  
D A E A E L A R V R E H M S H E V R D A

64621 TCCCGCTGGCCGCTGTTTCGACGTACGGACCCACCGCCTGGACGACGTCCGCACCCGGCTG 64680  
S R W P L F D V R T H R L D D V R T R L

64681 CACCTGAGCTTGGACCTGCTCATCGCCGACGCCACAGCGTCCACGTACTCACC GGCGAC 64740  
H L S L D L L I A D A H S V H V L T G D

64741 CTGCTCACCTTCTACGCCGACCCCGACGCGGCCCTGCGGCCCTCGGCTGCTCCTTCCGC 64800  
L L T F Y A D P D A A L P P L G C S F R

64801 GACTACGTCTTGGCCGTCCGCGCCACGCGGAGGCGAGCCGCGCCGCGCCCTCGAC 64860  
D Y V L A V R A H A E G E P R R R A L D

64861 CACTGGCGGGCCCGGCTGGCCGACCTGCGGGGCGCCCGGCTGCGGCTGCGGTGCCGG 64920  
H W R A R L A D L P G P P G L P L R C R

64921 CCCGAGGAGCTGACCGCGCCGCGGTTTCGCCCGCCTCACCACCGACTCGGCCCGACGCC 64980  
P E E L T A P R F A R L T T G L G P D A

64981 TGGGCACGGCTGCGGCGCGCCGCGGCGCGCCGAACTCACC CGCGCCGACTGATCTGC 65040  
W A R L R R A A A A E L T P A A L I C

65041 GCCGCCTTCTGCGACGTCTTCGCCAGTGGAGCGACACCCCGCTTCACCTCAACCTC 65100  
A A F C D V L A Q W S D T P R F T L N L

65101 ACCACCTTCCACCGCCCCCGCTGCTCCCGCGGTGGACGACCTCGTCGGCGACTTCACC 65160  
T T F H R P A L L P G V D D L V G D F T

65161 ACCACGACCTGCTCGGGGTCGACGCGAGGGGACACCTTCCGGGACCGGGCCCGCGA 65220  
T T T L L G V D G E G D T F R D R A R R

65221 CTCCAGGACCGCATCTGGGAGGACCTCGAACACCGCGTCGTCAGCGCGTCGAGGTCCTG 65280  
L Q D R I W E D L E H R V V S G V E V L

65281 CGGATGTGCGCCGCGAGCGGGGACCCACGACGCGTCCGGATGCCGGTCGTCTTCACC 65340  
R M L R R E R G T H D A V R M P V V F T

65341 AGCACCTGCGGGCCCGCGCCCGCGCCCGCGGACGCGCCCGCGCTGGCGGGTACGG 65400  
S T L R A A G P A P R T A P P A W R V R

65401 CCCGGCTACGCGATCAGCCAGACCCCGCAGGTCTGCTCGACCATCAGGTGAGCGAGAGC 65460  
P G Y A I S Q T P Q V L L D H Q V S E S

65461 GACGGCCGACTGGTCTGCACCTGGGACTACGTGCGGACGCCTACCCGCCCCGGGCTGATC 65520  
D G R L V C T W D Y V A D A Y P P G L I

65521 GAGGCCATGTTCTGGGGCCTTCGAGGCGCTCCTCGCTCGCTCGCCGTCACGACGACGAC 65580  
E A M F G A F E A L L A S L A G H D D D

65581 GCCGGCCACGACGACGACGCGCCGACGACGCGCCCGGCCACGACGACGCGCCCGGC 65640  
A G H D D D A G H D D G P G H D D G P G

65641 CACGACGACGGCCCCGGCCACGACGACGCGCCCGGCCACGACGACGCGCCCGCGAC 65700  
H D D G P G H D D G P G H D D G P G R D

65701 GACAGTGC CGATCACGGCCACAGTGCCACGCACGACGACGCGCCCGGAAACGACAGA 65760  
D S A D H G H S A T H D D S A A R N D R

65761 GAGGGAGGTGGACCGGAGTGACGAGCGCCCGGCCACGCGACACTGCTCCCCGCCGACC 65820  
E G G G P E \*  
M T S A R P T P T L L P A D Q (orf16)

65821 AGCGGGAGCTGCTGCGGATGATGAACGACCGCACCCGTCGCCCGCGCACACCTCA 65880  
R E L L R M M N D R T A P V P A H T L T

65881 CCGCCCAACTGGCCGACGCCGCGCGCACGCACGCGGGCTCTGGCACTGGTGGCACCGG 65940  
A Q L A D A A R T H D R A L A L V A P G



65941	GTCTGACACTGAGCCACGCCGAAGTGGACGCCCGGGCGCGCGGTGGCCGCCCGGCTCA L T L S H A E L D A R A A A V A A R L T	66000
66001	CCGCCGCGGGCGTCATCCCGGGGACCGGGTCGCCCTCGCCGTCGAGTACGGCTGGGAGC A A G V I P G D R V A L A V E Y G W E Q	66060
66061	AGGTCGTGGGCGCCCTGGCCGCGCTCCGCGCCGAGCCGTCTGCCTGCCCGTCGCCCCCG V V G A L A A L R A G A V C L P V A P G	66120
66121	GGCTGCCCCGGCCCCGCCCGTGGCAGCACGCCACCCGGGCCGGGGCGACGGCCGTCTCTCA L P R P A R W Q H A T R A G A T A V L T	66180
66181	CCCAGTCTTGCTCACCCAGCGCATCGACTGGCCGCGAGAACTGCCCGTCTCTCCGTGG Q S W L T Q R I D W P Q E L P V L S V D	66240
66241	ACGAACCCGGGCCGCCGGTACCAACCCACCACGCCCGGCCGACGGACGGTCCGCGACCG E P G P P V P P T T A P A D G R S A T D	66300
66301	ACGCCGCCTACCGGCTGGACGCCCCCGTCAGCCACCGCGCGATCACCACCGCCGCCCTGG A A Y R L D A P V S H R A I T T A A L E	66360
66361	AGATCGACCGCGCCTTCCGCGTCGGACCCGGCGACCGGCTCCTGGCCCTGGCCCCCGCG I D R A F R V G P G D R L L A L A P A D	66420
66421	ACTCGCGCTCGCTCTCTACGAACTGTTGGGGCCCTCCTGGCCGGTGGCGCCCTCGTCC S P L A L Y E L F G P L L A G A A L V L	66480
66481	TCACCCGGGACATCGACCTGCGCGATCCCGGAGCCCTGCACGAGGCGCTGCGCACCCACG T R D I D L R D P G A L H E A L R T H G	66540
66541	GCGTCACCTCTGGCACTCGCCGCCCCGCCCTCCTCGGCCTCCTCCTCGACCACCTCGCCG V T L W H S P P A L L G L L L D H L A D	66600
66601	ACCGGGGCGGCAAACTGCCCGAGTCGCTCCGGCTGGTGCTGCTCGGCGGCGAACGCCTCG R G G K L P E S L R L V L L G G E R L D	66660
66661	ACCCCGCCCTCGTCCGCGCGCTCCGCGAGAGCGCCCGCACCAGCCGGCCGTGCGCCACC P A L V R R V R E S A P H Q P A V A H L	66720
66721	TCTCCTCGGCCACCCCGTCCGGCCCCCTGGACCACCTGCCTGGAGACCGGCGACCTCGCCC S S A T P S G P W T T C L E T G D L A P	66780
66781	CGGAATGGCGCTCGGTCCCGTCCGCGCGCCCCCTGCCCAACCAGCGGGCGCACATCTGT E W R S V P V G A P L P N Q R A H I L S	66840
66841	CCGAGACCTTGGCGCCCTGCCCGGTCTGGGTACCGGCCGCTCCACTACGGCGGCGTCCG E T L R P C P V W V T G R L H Y G G V A	66900
66901	CCGCCGAGCCCCCACCAGGAGGAGCACGCACCCGCGACCGTCCCGCACCCGGAGACCG A E P P T G E E H A P A T V P H P E T G	66960
66961	GCGAACCGCTGCTGCGCACCGGGCTGTTCCGCCCGCTGCTGCCCGAGGGCCTGATCGACG E P L L R T G L F A R L L P E G L I D V	67020
67021	TCGTGCGCGACGAGACCGCCCGGATCAGCGTCCGCGACCGGCCCTGAACCTCCAGGACA V G D E T A R I S V R D R P L N L Q D T	67080
67081	CCGAGACCGCCCTCGCCGCCACGAGGACGTGCACTCCGCCGTGGTCTCCCGTCCGGGC E T A L A A H E D V H S A V V V P V G R	67140
67141	GGGGAGACGAGTCGCTCGCGCGGGTACGGCTCCACCCCGGCGCCACGGCCGGCCCCGACG G D E S L A R V R L H P G A T A G P D E	67200
67201	AACTCCTCGCCCATCTGCGCCGCAAGGTCTCCCTTACCTGCTGCCCGGCCACATCGAGG L L A H L R R K V S P Y L L P G H I E V	67260
67261	TGGGCGGTCCGCTGCCGCTCACCCGGGACGGGCGGTGGACCGCGCGCGTCCACCGCCG G G P L P L T R D G R V D R A R V T A E	67320
67321	AGGCCCCCGCCCCGCTGCCGTGCCCGCCCGCGCGCGGCGGCGTCCGCACCCGCGCGGG A P A P A A V P A A A P A A S A P A R D	67380
67381	ACGAGGCCGAATCCTCGCCCAAGTGGCCCGGGTGACCTGCCGGGTGCTGGGAATCGGCG E A E L L A Q V A R V T C R V L G I G A	67440

67441	CCGTCGAACCCGATATGAACCTGCTCGACGCCGGTGCCACCTCCGTGCAACTCGTCCGCC V E P D M N L L D A G A T S V E L V R L	67500
67501	TGGCGACCGCTCTGGAGGAGGAACTCGGCCTCGACACCGACATCGAGGAACTGCTGGCCT A T A L E E E L G L D T D I E E L L A F	67560
67561	TCCCGTCGGTCGCCGTGATCGTCGGCCGCCACCTCGGCCGCCGGACGGCACCCACGGCCC P S V A V I V G R H L G R R T A P P A R	67620
67621	GGGACCCCTGCGCCCCGCGTCCGTAGCGTTTCGACCCGGGTCCGTACTGCCCGCGCCGC D P L P P A S V A F A P G S V L P A P P	67680
67681	CCGCGCCCGGACCCGTGCCGCCCGCGTCCGTGCCGCCCGCACCCGCGTCCGTACC GCCCG A P G P V P P A S V P P A P A S V P P A	67740
67741	CGTCCGAGTCTCTACCGCTCGCGCCGCCCGCACCCGGGCCCGTGCCACCCACGCCCCGTCC S E S S P L A P P A P G P V P P T P V P	67800
67801	CGCCCGCCTCCGTCCCGCCCCGCGTCCGGGGCCGCGCCGCACGTACCGCCCGCGCCGCCCG P A S V P P A S G A A P H V P P A P P A	67860
67861	CACCCATCCCCGCGCCCTCCGTGCCccccgcgccccgccccaaccggccctgtcaccg P I P A P S V P P A P R P Q P P L L T G	67920
67921	gcacggcgccccgccaggcgtTCAAGGACGCCACACGGCATCCGGCACGAGTTTCGACG I G A R Q A F K D A H H G I R H E F D A	67980
67981	CCACCGACGGCGTCGCCCTCAGCGGCCCGGACGACCACCACTACCGCCCGTCGCAGCC T D G V A L S G P D D H H L T A R R S H	68040
68041	ACCACCGCTTCGACCCCGGCCCGGTGACGCTGCCCGACCTGGCCGCCCTCTCTGGGGCCC H R F D P G P V T L P D L A A L L G A L	68100
68101	TCCGCCGGGTCCGCGGCCCGGGAGGCGAACCCAAATACGCCTATCCGTCCGCCGGTTCTCT R R V R G P G G E P K Y A Y P S A G S S	68160
68161	CCTACCCCGTCCAGACCTACCTGCTCGTCCACCCGGGGAAGGTGACCGGACTGCCCGGCG Y P V Q T Y L L V H P G K V T G L P G G	68220
68221	GCAGCCACTACGTCCACCCCGCGCGCAACCGCCTGGTGAGCATCGACCCACCGCGACCC S H Y V H P A R N R L V S I D P T A T L	68280
68281	TGCCCCCGACGCGCACGCCGAGATCAACCGCGCCGCTACGGGGAGGCGGCCTTCTCCC P A D A H A E I N R A A Y G E A A F S L	68340
68341	TCTACCTCATCGCCGCGATCGACGCGATCACACCGCTCTACGGCGATCTCTCTGGGACT Y L I A A I D A I T P L Y G D L S W D F	68400
68401	TCACCGTCTTCGAGGCCGGTGCCATGACCCAGTTGCTGATGCGGACCGCGCTCGGCACCG T V F E A G A M T Q L L M R T A V G T G	68460
68461	GCATCGGCCTGTGCCCCGTGCGCACGATGGACCCCGCGCCGCTGCGCCGCGCGTTCCGCC I G L C P V G T M D P A P L R R A F A L	68520
68521	TCACCGACCGGCACCGCTTTCGTCCACGCCCTCTCGGCGGGCGGCCCGCACGGAGGCC T D R H R F V H A L L G G R P R T E A P	68580
68581	CGTGAACCGGCACGGCCCCCTGGCGGGCCGGCGGAGAGCGTCGACACCCGACGCGCCGC M N R H G P L A G R R Q S V D T R S A A *	68640 (orf15)
68641	GTGGGTGGCGCCGACGGGCACCCGGGGCTGCCGCTGGAGGTGGCCGCCACCCGGGACGG W V A P T G T P G L P L E V A A T R D G	68700
68701	CGTCGACCCGGCCGAATGGGCCCCGACCCACCTCGACACCGTCACCGGCTGGCTGCACCG V D P A E W A R T H L D T V T G W L H R	68760
68761	TCACGGAGCCGTCTGTTCGCGGCTTCGGCGTCGGCCTCGACGGCTTCGCGACGTCGT H G A V L F R G F G V G L D G F G D V V	68820
68821	CCACGCCCTGGCCGGATCCCCGAGGCGTAGCTCGAACGGTTCGTGCGCGCGCACCCGCCCT H A L A G S P E A Y V E R S S P R T A L	68880

68881 CGGGCATCACCTCTACACGCCACCGACCACCCGCCGACCAGCCCATCCCCCGCACAA 68940  
G H H L Y T A T D H P A D Q P I P P H N

68941 CGAGAACTCCTACCAACTCCGCTTCCCCGGACGGCTGGTCTTCGGCTGCCTCACCCTGGC 69000  
E N S Y Q L R F P G R L V F G C L T P A

69001 CCGGACCGGGCGGCGGACCCCGCTCGCCGACACCCGGCGCGTCTGGGCGCCTCGACCC 69060  
R T G G A T P L A D T R R V L G R L D P

69061 CGCCCTCGTCGCGCCCTTCGCCCCGCCGGGGTGCTCTACCAGCGCAACTACGGCGACGG 69120  
A L V A A F A R R G V L Y Q R N Y G D G

69121 GATCGGCATGTCTTGGCAGGACGCCTTCCAGACCCGCGACAAGGCGGCCGTCACCGCCTA 69180  
I G M S W Q D A F Q T R D K A A V T A Y

69181 CTGCGCCCGCCCGCGCTCGACGTGGAATGGAAACCCGACGGCGGGCTGCGGACCACCCA 69240  
C A A R R V D V E W K P D G G L R T T Q

69241 GGTCCGCCCCCGCCCTCGCGCTCCACCCGGCGACGGGGGAGCGGGTGTTCAACCACGC 69300  
V R P A L A V H P A T G E R V W F N H A

69301 CGCGTTCTTCCACGTCTCCGCCCCGGCCCGCGCTGCGGGACGCCCTGCTGGCCCACTT 69360  
A F F H V S A R P P A L R D A L L A Q F

69361 CGACGAACGCGACCTGCCGAGCCACTCCTGTACGGCGACGGCCGGCCCATCGAACCCTGC 69420  
D E R D L P S H S C Y G D G R P I E P A

69421 CGTCATGGAGGAACTGCACCACGCCTACGCCGCCGAACTGGTGGCGCCCGCTGGCGGGC 69480  
V M E E L H H A Y A A E L V A P A W R A

69481 CGGCGACGTCTCTCTCGTCGACAACCTCCTCACCGCGCACGGCAGGGAACCTTCACCGG 69540  
G D V L L V D N L L T A H G R E P F T G

69541 CGAACGCCGCGTCTGTCGTCGGCATGGCACAGCCGCTGGACTGGGACGAGGTGAGCGCGTG 69600  
E R R V V V G M A Q P L D W D E V S A \*  
M (orf14)

69601 ACCGCCCCCGGCACACCGCTGCCCGCGACCTTCGTCCAGCGCGGCTGTGGCCGTCCACT 69660  
T A P G T P L P A T F V Q R G L W P S T

69661 CGCCACGCCCCCGCGGAGGTCACCCACGTCCGCGCCCTGCGCCTGACCGGGGACACC 69720  
R H A R P A E V T H V R A L R L T G D T

69721 GACACGGCGCGGCTCACCGAGGCGCTCCGGCGGGTCACCGCCGCCCTCCCCGCCCTCACC 69780  
D T A R L T E A V R R V T A A L P A L T

69781 GCCGAACTCTCCGGCGACGAGGAACCCCGCCTGACCCTCCGGCCGGACGCCCCGAGGTC 69840  
A E L S G D E E P R L T L R P D A P E V

69841 ACCCCGGTCTGACCTGCGCGGAGCCCCGTCCGCCGGACGCGACGCCGTCTGCGTGGCGCTG 69900  
T P V D L R G A P S A G R D A V C V A L

69901 CTGCGCGCCGACCGGGACCCCTCGCGCCGGACGCCACCGGGCCCGCTTCCACCTGGTG 69960  
L R A D R D H P R A G R H R A R F H L V

69961 CGGCTCCACGACGACGAGACGGTGCTCGCGCTCACGGCCACACCTCCTCCTCGACACA 70020  
R L H D D E T V L A L T A H T L L L D T

70021 CCGTCTCTCTACGCCGTGCTCGGCGGGTCTGCCAGGCGTACGCCGGCCGCTTCCGCCCC 70080  
P S L Y A V L G A V C Q A Y A G R F R P

70081 GAGCACTACCGCGACGCCACCCCTGCCCGACGCGCCCCACGCCCCCTCTCCGGTCCG 70140  
E H Y R D A T T L P D A P H A P L S G R

70141 GCCCGGGCCTCCCGCCGGCGCTGGTGGCACCGGCGCCTGGCCGCCCTGCCCGGCCGGCC 70200  
A R A S R R R W W H R R L A A L P G P A

70201 CCGGCCCCCGCCGGCCCCCGCGACCGGGTGACCGAAACCCACGGCTGCGCATCCCC 70260  
P A P A G P P R D R V T E T H R L R I P

70261 GCAGCGCGCTGGAAAGCCCTGACCGCCCTGACCGCCCTGGGCGGCCCTCGGCGGCAAC 70320  
A A R W K A L T A L T A L G G P L G G N

70321 GGCTCGCTCGCCGTATGGCCCTGGCCGCTGGTGCCTGCGGCCCCGGACACCGGGGA 70380

G S L A V M A L A A W C L R A P D H R G  
 70381 CCGGCCCGCTTACACACCGTCGTGACCTGCGCGACACCTCGGACTCGGGCCCGCCGTC 70440  
 P A R F T T V V D L R D H L G L G P A V  
 70441 GGCCCGTTCACCGACCGCCTCGTCTTCGGCGCCGACCTCGGCGAAGCGCCGCGCCCTCC 70500  
 G P F T D R L V F G A D L G E A P R P S  
 70501 TTCCGGGACGTACGCTGCGCGCCAGTCCGGGTTCCTGGACGCCGTCGTGCACTACCTC 70560  
 F R D V T L R A Q S G F L D A V V H Y L  
 70561 CCCTACGGCGACGTGCTGGAACTCGCGAGGAACTGGGCCGCGTCACCGCGCCCCGACCC 70620  
 P Y G D V V E L G R E L G R V T A P R T  
 70621 GCCGCGCACTGGGACGTGGCGCTGAACCTTCTGCCGCAACCCGCCACCAGCGCCGCCACC 70680  
 A A H W D V A L N F C R N P P T S A A T  
 70681 CGCGGCGAACGCACCCTCGCCGAACGCGGCCTGTCCATCGAGCTGTTCCGCGAGGCCGAC 70740  
 R G E R T L A E R G L S I E L F R E A D  
 70741 CTGCTCGGCGCGGCCCGCACCGGTCCCGCGCACCGGTGGGACGGCACGGTGTCTCGCCCTC 70800  
 L L G A A G T G P A H R W D G T V L A L  
 70801 TCCCTAGGCGAACTCGGCGACGACACCGTGTGCTCCTCGACGCCGACCGCGACACCCG 70860  
 S L G E L G D D T V L V L D A D R D H P  
 70861 CACCACGGAACCGCCGACCGGTGCTCCACCGGATGGACGAAGCGCTCCTGGCGGCCGTC 70920  
 H H G T A D R L L H R M D E A L L A A V  
 70921 GCCGACCCGGACGCCCCCTGCCCCCTTGCCCGCCCCCGCGCACACCAAGAGGCCAC 70980  
 A D P D A P L P P L P A P A H T T R S H  
 70981 CGATGACCACGACCCCGCGGACCGCCGCCGAGCCACCTACCACGTGGTGGTCAACGACG 71040  
 M T T T P R T A A E P T Y H V V V N D E (orf13)  
 R \*  
 71041 AGGAGCAGTACTCGATCTGGCTCGCCGAACAGGAGATCCCGCCGGCTGGCGGGCCACCG 71100  
 E Q Y S I W L A E Q E I P A G W R A T G  
 71101 GAACCTCCGGCACCCAGGAGGAGTGCCTGCGCCACATCGACGAGGTGTGGACCGACATGC 71160  
 T S G T Q E E C L R H I D E V W T D M R  
 71161 GCCCCCGCAGCCTGCGCGAGGCCATGGCCGCGCGGAGCACGCGGAGCCCGCTCCCGCCC 71220  
 P R S L R E A M A A A E H A E P A P A P  
 71221 CGGCCCCGCGCAGGAGGAGCCGAGCCTCGTCGACCGGCTCTGCGCGGGCGACCAGCCGG 71280  
 A P A E E E P S L V D R L C A G D Q P V  
 71281 TGGAGTCGGTCTCTCCGCCCGGAGCGCACGGCCCGCCCTGCGGGAGGCCGTCGACCGCG 71340  
 E S V L R P E R T A A A L R E A V D R G  
 71341 GCTACGTCTTCGTCCGCTTCGCGCCACCCGCGCGGCACCGAACTCGGCGTCGCCGTCG 71400  
 Y V F V R F A A T R G G T E L G V A V D  
 71401 ACCCCGCGGCGACCACTGGACGGCACCGAGTGTGCGCCTGACCGGCACCCCTACCCCTCG 71460  
 P A A T T M D G T E L R L T G T L T L D  
 71461 ACTTCGAACCGGTCCGCTGCCACGCCCCGCTCGACGTGACCACCTTCACGGGCGAGGGCC 71520  
 F E P V R C H A R V D V T T F T G E G R  
 71521 GCCTGGAGCGCGTGTCCGGCACCTGACCCCCCGCGCCACCCGGCCGTGAGGCGCGGCTC 71580  
 L E R V S G T \*  
 71581 GGGACCGGGCCGCGACCCACCGAAGGGAGGGACCCCATGACCACCCCATGACCACCCC 71640  
 M T T P M T T P (orf12)  
 71641 CACGACCACCCGACCAACACCCGACCGCCGCTCTTCGCCACCTCCGCGCCCCCGGCTC 71700  
 T T T R T T T R T A V F A H L R A P G L  
 71701 CGGCGACCTCTCCAGCGCAACATCGGCCTCGCCCTCGTCCCGCGCCCGCCGCGGCGAC 71760  
 G D L L Q R N I G L A L V R R A R P A T  
 71761 GGCGGTACCCCTGGTCTCGGCGAGGACCTGGCGGCCCGCTTCGGTCCGGCACTCACCCG 71820  
 A V T L V V G E D L A A R F G P A L T R

71821 CCACACGTACGCCACCGACGTGCTGCCCTGCCCCAGCGGGGCGAGGCCGACCCCCGGTG 71880  
H T Y A T D V L P C P Q R G E A D P R W

71881 GCCCGCCTTCTGCGCACCTGCGCCGACCGCGCTTCGCCCTCGCCGTCGTCGACCCGGA 71940  
P A F L R T L A D R R F A L A V V D P D

71941 CAGCCAGGGCCTGCACGCCGCCACGCCCGGGCCGCCGCGTGCCTCGAGCGGATCGGCCT 72000  
S Q G L H A G H A R A A G V P E R I G L

72001 GCCGCAGGACCGGCCCGGAGACGAACACATCACCCATCCCATCCGCTCCACGTCCCCCT 72060  
P Q D R P G D E H I T H P I R L P R P L

72061 GTGGGGGACCCCGACCTGTACGAGTACGCCACTGCCCTCGCCCGCGCTGGGCCTGCC 72120  
W G T P D L Y E Y A T A L A A A L G L P

72121 CGCACCGCGCGCCCCGGCGACGTCTGCGGAGCTGCCCCGCACCCGCGGCGTCCGCC 72180  
A P P R P G D V L P E L P R T R G V R P

72181 GCCGACGGCCGGTCTGCCCCGTCCGCTCGTCGCGCTCCACCCCGCGGGGCACCGCACTG 72240  
P T A G L P R P L V A V H P G G A P H W

72241 GAACAGGAGATGGCCGCTCGAGCACTACGCCCGGCTCTGCGCCCGCTCGCGGCCGAAC 72300  
N R R W P L E H Y A R L C A R L A A E L

72301 CTCGGCCTCCCTCTGCCTGCTGGGCGACGAAGCCGAACGCCCGGAGCTGGAAGTCTCCG 72360  
S A S L C L L G D E A E R P E L E L L R

72361 GCACGCCGCTCTGACCGCGTCCCCGCGAGCCGTCGTCCACCTCGAGGCGGGCGCGGACCT 72420  
H A V L T R S P R A V V H L E A G A D L

72421 CGACCGGACCGGAACGTCTCGCGACGCCGACCTGCTCGTCGGCAACGACTCCTCGCT 72480  
D R T A N V L A D A D L L V G N D S S L

72481 CGCCACGTGCGCGCCGCCCTCCGCACCCCGTCCGTGCTCTCTACGGCCCGACCGGCAC 72540  
A H V A A A V R T P S V V L Y G P T G T

72541 CGAGTACCTGTGGACAGGATCTACCCGTACCACCGCGGGGTCTCCCTGCGGTGGCCGTG 72600  
E Y L W T R I Y P Y H R G V S L R W P C

72601 CCAGCGGCTGCGGCACGCCCGAGGCGAACTCGCCGGCCGGCGGTGCGCGCACGGCTGCGT 72660  
Q R L R H A A G E L A G R R C A H G C V

72661 CCTGCCCTACCAGGGCCCGGCCCGCCCGTATCCGCGCTGTCTGGCCGACCTGCCGTTGA 72720  
L P Y Q G P A G P Y P R C L A D L P V D

72721 CAGGGTCTGGCCGGCGGTGACCGCCCGATGGGCGAGCCCCACCCCGTGACGATCAGGAG 72780  
R V W P A V T A R W A S P H P V T I R S

72781 TACCCCATGAGCGCCGACCCGTCCCGGGTGGCGACGATCCTCTCCGTCAACTTCAACCAC 72840  
T P \*  
M S A D P S R V R T I L S V N F N H (orf11)

72841 GACGGCTCCGGCGTGCTGTTGCGGGAGGGCAGGATCGCCGGCTACGTCAACACCGAGCGC 72900  
D G S G V L L R E G R I A G Y V T T E R

72901 CGCTCCCGCCTCAAGAAGCACCCGGGCGTGCAGGAGGACCTCGACGAAGTCTGGAC 72960  
R S R L K K H P G L R E E D L D E L L D

72961 CAGGCCGGGGCCGACCTCTCCGACATCGACCACGTATGCTCTGCAACCTGCACACCATG 73020  
Q A G A D L S D I D H V M L C N L H T M

73021 GACACACCCGACATACCCCGGCTGCACGGCTCCGACCTCAAGGAGACCTGGCTCGCGTTC 73080  
D T P D I P R L H G S D L K E T W L A F

73081 TGGGTCAACCAGCGCAACGACGAGGTGAGCCTGCGCGGCCCGCGCATCCCTGCACCGTC 73140  
W V N Q R N D E V S L R G R R I P C T V

73141 AACCCGGACCAACCTCATCCACGCCGCCACCGCCTACTACACCTCCGGCTACGACTCG 73200  
N P D H H L I H A A T A Y Y T S G Y D S

73201 GCGATGGCCGTGGCCATCGACCCACCGGCTGCCGCGCCTTCGCCGGAAGGGCAGCCGC 73260  
A M A V A I D P T G C R A F A G K G S R

73261 CTCTACCCCTGCGCCGCGACCTCGACGCCTGGTTCAACGCCAACATCGGCTACTGCTAC 73320  
L Y P L R R D L D A W F N A N I G Y C Y

73321 GTCGCCGACCTGATGTTCCGGCTCCAGCATCGTCGGCGCCGGCAAGGTCATGGGCCTCGCC 73380  
V A D L M F G S S I V G A G K V M G L A

73381 CCCTACGGCAGACCCGCCGACGGCGCCGCCCGACGAGGAACCGCCGAGACCGTGC GC 73440  
P Y G R P A D G A G P D E E P P E T V R

73441 GACTTCGCCGCCCTGGTGGCCCTGGCCGACCGGCACCCGCGCCTCGTCGACGTCGACGGC 73500  
D F A A L V A L A D R H P R L V D V D G

73501 AGGAAGCTCAACGCCACCCTCGCCCACTACATCCAGCTGGGCCTGGAACGCCAGCTGACC 73560  
R K L N A T L A H Y I Q L G L E R Q L T

73561 GCCGTCTTCGCCGAGCTCGCCCCGCTGTGCGCCCGCAACGGCATCGCACCCGACATCTGC 73620  
A V F A E L A P L C A R N G I A P D I C

73621 CTCTCCGGCGGTACCGCCCTCAACGCCATCGCCACCCAACTCGCCTTCGAGTCGACCGGC 73680  
L S G G T A L N A I A T Q L A F E S T G

73681 TTCGAGCGCATGCACCTCCACCCCGCCTGCGGCGACGACGGCACCGCATCGGCGCGGCG 73740  
F E R M H L H P A C G D D G T A I G A A

73741 CTCTGGCACTGGCACCACGTCTGGGCCACCCCGGCTCCACCACACCAACGCCGACCTC 73800  
L W H W H H V L G H P R L H H T N A D L

73801 ATGTACTCCGTCCGTGAGTACCCCGAGCACACCGTCCGGCGGGCCGTGCGGGACCACGCG 73860  
M Y S V R E Y P E H T V R R A V R D H A

73861 GCCGACCTCGTCGTCGAGGAGACCGGCGACTACGTGCCAGGGCCGCCGAAGTGGTCGCC 73920  
A D L V V E E T G D Y V A R A A E L V A

73921 GGCGGCGCCGTATCGGCTGGTACGACGGCGCCGGCGAGGTGCGGGCCGCGGGCCCTGGGC 73980  
G G A V I G W Y D G A G E V G P R A L G

73981 CACCGCAGCATCGTCGCCGACCCGCGCGACCCCGCATGCGGGACCGGCTCAACTCCCAG 74040  
H R S I V A D P R D P A M R D R L N S Q

74041 GTCAAGTTCGCGAACACTTCCGGCCCTTCGCGCCGTCCGTGCTCAAGGAGCACGCCGCG 74100  
V K F R E H F R P F A P S V L K E H A A

74101 GAGTGGTTCGGCCTCTCCGACAGCCCCTTCATGTGCGGGCCACCCCGTCTCAAGCCC 74160  
E W F G L S D S P F M L R A T P V L K P

74161 GGCGTGCCCGCATCACCCACGTGACGGGACGTGAGGATCCAGTCGGTCACCCGCCAG 74220  
G V P A I T H V D G T S R I Q S V T R Q

74221 GACACCCCGCCTTCCACGACCTCATCCACGCCTTCAAGGACCGTACGGGGATCCCCATG 74280  
D T P A F H D L I H A F K D R T G I P M

74281 GTGCTCAACACCAGCCTCAACACCAAGGGCGAGCCGATCGCGGAGACACCCGAGGACGCC 74340  
V L N T S L N T K G E P I A E T P E D A

74341 CTGCGCACCTGCTCGGCTCCCGGCTCGACCACCTGGTGCTCCCGGGCCTCATCGTCAGC 74400  
L R T L L G S R L D H L V L P G L I V S

74401 GGCCGGACGGCGGCCCGCTCATGAGCGCCCGCGGGGCGAGCGGACCCGGCGCCGCGCGC 74460  
M S A P R G E R T R R R A L (orf10)  
G R T A A R S \*

74461 TCGAACCGACATCGCCGCGATCTGGGCCGAGACCCTCGGCAGGGACAGCGTCGGCCCGC 74520  
E R D I A A I W A E T L G R D S V G P H

74521 ACGAGGACTTCGCCGCGCTGGGCGGCAACTCCATCCACGCCATCAAGATCACCAACCGGG 74580  
E D F A A L G G N S I H A I K I T N R V

74581 TGGAGGAACTCGTCGACGCCGAGCTGTCCATCCGCGTCTCTGCTCGAGACGCGCACCGTGG 74640  
E E L V D A E L S I R V L L E T R T V A

74641 CCGGCATGACGGACACGTCCACGCCACGCTCACGGGGGAGCGGGACCGGTGAACACCGA 74700  
G M T D H V H A T L T G E R D R \*  
M N T D (orf9)

74701	CCTGCCCCGGCTGCTCGACCGGATCGCCGGCCTGCGCGTCTCGTCATCGGCGACGTCAT	74760
	L P R L L D R I A G L R V L V I G D V I	
74761	CCTCGACACCTACGTCTGGGGAGCCACCTCGGGCCTGTGCCGGAATCCCCCGTCCCTGC	74820
	L D T Y V W G A T S G L C R E S P V P A	
74821	CGTCACCCTGACCTCCGTGCGCCACAGTGCGGCGCGCCGCAACGTGCGCGTGAACCT	74880
	V T L T S V A H Q C G G A A N V A V N L	
74881	CCGGGCGCTCGGCGCCGAACCGGTGCTGCTCTCCGCGACGGGTGACGACCGCGCCGGCCG	74940
	R A L G A E P V L L S A T G D D R A G R	
74941	CCGGCTGCGCGAAGCCCTCCGTGCGCGGGACGTGACACCGGCGGACTCTTCGTACAGCC	75000
	R L R E A L R A R D V D T G G L F V Q P	
75001	CGGCCGGACACGGTCACCAAACGCCGCGTCATGGCCGACGGACAGATGCTGCTCCGCCT	75060
	G R T T V T K R R V M A D G Q M L L R L	
75061	CGACGAGGGCGGCGAACACCCGTTGCCCGTGGCGACGGACACCGGAAGCCGCTGCTCGA	75120
	D E G G E H P L P V A T D T G S R L L E	
75121	ACGGGCCGCCGGCCTGCTGCCCCGCGTCGACGCCGTGATCGTCTCCGACTACGGGTACGG	75180
	R A A G L L P A V D A V I V S D Y G Y G	
75181	CGTGTGGGAGCCCCGACACCGTCCGCCGCGCTCGCCGCACACCGCGAACTCGGCCCGTCCAC	75240
	V W E P D T V A R L A A H R E L G P S T	
75241	CCTGGTCGTCGACTCCCGCCGGCCCGCGCGCTTACCGCGCTGCGGGCCAGCGCCGTCAA	75300
	L V V D S R R P A R F T A L R A S A V K	
75301	ACCCAACCACGCGGAGGCGCTGCGCCTGCTCGACGCCGGCGAACCCCCGCCCGGCCGGC	75360
	P N H A E A L R L L D A G E P P P G P A	
75361	CAGGGCGGACTGGGCGGCCCGCCTCGGCGACCGGCTCCTGCGCCTGACGGGAGCCGAACG	75420
	R A D W A A A L G D R L L R L T G A E R	
75421	GGTCGCCCTCACCCCTGGACGCCGACGATCACTGTCTTGAACGCGACCGGCCCGCGGT	75480
	V A L T L D A D G S L L F E R D R P P V	
75481	CCGCACGTTGCCCCGGGGCAGCCGGGCACCGGTACGGCCGCCGTGCGCGCCGGCGACGC	75540
	R T F A R G S R A P V T A A V G A G D A	
75541	CTTACCGCGGCCCTCACCCCTCGCCCTCGCCGCGCGCGCGACTCCGCGGTGCGCGCGGA	75600
	F T A A L T L A L A A G A D S A V A A E	
75601	ACTGGCCTCCGCCCGCCCGGCACGGCCGTGCGCACCCCCGGCACCAGCACCTGGCACGC	75660
	L A S A A A G T A V A T P G T S T W H A	
75661	CGACGAACTGCGCCGACTGCTCGGCGGCACCGGCAAGGTCTGCCGGACCGGCACCCTGCC	75720
	D E L R R L L G G T G K V C R T G T L P	
75721	CGCCCGGCTGCTCGACCCGGCCCGCCGCGACCGCGGGTCTCTTACCAACGGCTGCTT	75780
	A R L L D P A A R D R R V V F T N G C F	
75781	CGACCTCCTGCACGGCGGCCACGTCTCCTGCCTGAGCCGGGCAAGGAACTGGGCGACCT	75840
	D L L H G G H V S C L S R A K E L G D L	
75841	GCTCGTCGTCGGCGTCAACTCCGACGCGAGCGTCCGACGCCTCAAGGGCCCCCGTCGCC	75900
	L V V G V N S D A S V R R L K G P R R P	
75901	GGTGATCCCCCTCGCCGAACGCATGCGCGTCTCGCCGCCCTGAGCTGCGTGGACCTCGT	75960
	V I P L A E R M R V L A A L S C V D L V	
75961	CGTGCCTTTCGACGACGACAGCCCCGCCGCCCTCATCGAGGCCCTCCGCCCCGAGGTCTA	76020
	V P F D D D S P A A L I E A L R P E V Y	
76021	CGCCAAGGGCGGGGACTACCCCTCGCGACCCTGCCCCGAAGCACCCCTCGTCCAACGGCT	76080
	A K G G D Y T L A T L P E A P L V Q R L	
76081	CGGCGGCGTCTCCACCTGCTCCCCAGCGTCGCCGACACCTCCACCACCGACATCATCCG	76140
	G G V V H L L P S V A D T S T T D I I R	
76141	GCGCATCCACGCCCTGTCCAGGACCGGCGAGGGAGACACCCCATGAGCCACGCCATCGGA	76200
	M S H A I G (orf8)	

R I H A L S R T G E G D T P \*

76201	CCGAGCCGGCTGATCCCCGCCATCCGCGAAGCGCTCGGGGACGAGAAGGACCCCCGGCTC	76260
	P S R L I P A I R E A L G D E K D P R L	
76261	GCCCTCTACGTCCACGTCCCCTTCTGCTCCTCCAAGTGCCACTTCTGCGACTGGGTCACC	76320
	A L Y V H V P F C S S K C H F C D W V T	
76321	GACATCCCCGTGCGACGCCTGCGCGGCGACAGCCGGAACGCTCGCCCTACGTACCCGCC	76380
	D I P V A R L R G D S R E R S P Y V T A	
76381	CTCTGCGACCAGATCCGCTTCTACGGCCCCCAGCTCACCCGGCTCGGCTACCGCCCCGAG	76440
	L C D Q I R F Y G P Q L T R L G Y R P E	
76441	GTCATGTACTGGGGCGGGCGCACCCCCACCCGGCTCACCGGCGACGAGATGACGGCCGTC	76500
	V M Y W G G G T P T R L T G D E M T A V	
76501	CACCAGGCCCTCGACGACGCCTTCGACCTGACGGGACTCCGCCAGTGGTCGGTGGAGAGC	76560
	H Q A L D D A F D L T G L R Q W S V E S	
76561	ACCCCGAACGACCTCGACCCCGCCACCCCTCGACACCCCTGCGCGGCCTCGGCGTCACCCGC	76620
	T P N D L D P A T L D T L R G L G V T R	
76621	GTCAGCGTCGGCGTCCAGTCGCTCAACCCGTACCAGCTGCGCAAGGCAGGCCGGGCCAC	76680
	V S V G V Q S L N P Y Q L R K A G R A H	
76681	TCGCGCGAACAGGCCCTGGCGCGCGTCCCCCTGTTGCGCGCGCGCCGCATCGACGAGTTC	76740
	S R E Q A L A A V P L L R R A G I D E F	
76741	AACGTCGACCTGATCGCCGGCTTCCCCGGCGAAGCCGTGAGTCCTTCGAGGAGACCCTG	76800
	N V D L I A G F P G E A V E S F E E T L	
76801	CGCACCGTCCTCGCGCTCGACCCCGCCGACGTCTCCGTCTACCCCTACCGCGCCACCCCC	76860
	R T V L A L D P P H V S V Y P Y R A T P	
76861	AAGACGGTCATGGCCATGCAGCTCGACCGCGAGTTCGTCGAGGCCCGGAACCGGGACGGC	76920
	K T V M A M Q L D R E F V E A R N R D G	
76921	ATGATCGACGCCTATGAACGGGCCATGGCCGCGCTCGGCGCCGCGGCTATCACGAGTAC	76980
	M I D A Y E R A M A A L G A A G Y H E Y	
76981	TGCCACGGCTACTGGGTGCGGACGCGCGCCACGAGGACCAGGACGGCAACTACAAGTAC	77040
	C H G Y W V R D A R H E D Q D G N Y K Y	
77041	GACCTGGCCGGCGACAAGATCGGCTTTGGCAGCGGCGCCGAATCGATCATCGGTCACCAC	77100
	D L A G D K I G F G S G A E S I I G H H	
77101	CTGCTCTGGAACGAGAACAGCGCCTACGCCCCGTACCTGCTCGCCCCCGCGAGTTCTCC	77160
	L L W N E N S A Y A R Y L L A P R E F S	
77161	GCCGCCCACCGGTTTACCACCGCCGAACCCGACCGCCTGACCGCCCCCGTCGGCGGCGCG	77220
	A A H R F T T A E P D R L T A P V G G A	
77221	CTGATGACCCGTGAAGGCGTGGTCTTCGCCCCGTTCGCGAGACTGACCGGCCTGGACTTC	77280
	L M T R E G V V F A R F R R L T G L D F	
77281	GCGGACGTCCGCGCCACACCGTACTTCCGCCAGTGGTTCGAGCTCCTGGAGCGCTGCGGC	77340
	A D V R A T P Y F R Q W F E L L E R C G	
77341	GGCCGCTTCGTCGAGACGCCGTACAGCCTCCGCTGGAGCCGTCCACCATCCACCGCGCC	77400
	G R F V E T P Y S L R L E P S T I H R A	
77401	TACATCACCCACCTCGCCTACACCATGGCCCATGGCCTGGCCCCCGAACGCGCCTGA	77457
	Y I T H L A Y T M A H G L A P E R A *	

# SEQ ID NO: 2 ORFs BLM gene cluster ORFs 31-40

(notice this part is on the reverse strand and the last nucleotide (18660) is the first (1) on the whole cluster of 77457 bp. Also the last orf (40) is incomplete and contains frame shifts)



1 GTGACCGAGAACCTTCCGTCGTGCCCCGAATGCTCCAGCGGTACACCTATGAGATGGGT 60  
M T E N L P S C P E C S S A Y T Y E M G  
(orf31)

61 GCGCTCCTGGTCTGCCCCGAATGCGGCCACGAGTGGCCGCCGCGACCGCCGAGTCCGCG 120  
A L L V C P E C G H E W P P A T A E S A

121 GACAACCCCGAAGACGGCGGATCAGGGACGCGGTGCGCAACGTA CTGCGCGACGGCGAC 180  
D N P E D G A I R D A V G N V L A D G D

181 ACCGTCACGGTGGTCAAGAGCCTGAAGGTCAAGGGCCACCCGACCGGCATCAAGGCCGGC 240  
T V T V V K S L K V K G H P T G I K A G

241 ACCAAGGTGCGCAACATCCGCCTCGTGGAGGGTGTGGCCGGCCACGACATCGACTGCAAG 300  
T K V R N I R L V E G V A G H D I D C K

301 ATCGACGGGTTCGGCGCCATGCAGCTCAAGTCCAGCGTGGTCAAGAAGGTCTGACCGGTT 360  
I D G F G A M Q L K S S V V K K V \*

361 ACGCCGGCCAGGCCCTGCCAGGCTCCACTACGCCGCGCGCAACCGAGCCGGAACGGG 420

421 GCCCGGGCCCGCTCCAAGTCCCGTTCGTCGCGGGCCGCGCAGCCAGGCCGTGTTACC 480

481 CTGGGGTCGCCGTCCCGTTTCGCACGCGTCTACACGCCACCACGCACGGCACGGAACCTC 540

541 CCCGAACCTGCGCACGTTCCCAAGTCCCGCGTGCCCGGATCCGCCCGACCGGCGTCGG 600

601 TCCGCCCGCCGGGCGCGGGCCGGTCCCGGGCCGCGGCGGAGGGGGTCTCGCGCCGTG 660

661 GAACGCCGCGCGGAAATTTACGTATAGGTAGAGATCCCGGCGAAGCGATCGGCGCGTTAT 720

721 GGCAGCATCCGCGCGGGCCCGCGCGCAGTTCTCGGTCCCGGACCGATGGCGTCAAAAG 780

781 TGAGCGACGAAATCGCCGGATCGCGCGAGGACCGTCGCGGGCCGCGACGAGACAACCGGG 840

841 GGATATATCAGCGCATTCCAGGTCACGCGTTGACTGGAAATCGCCTACTTATCGCGTCA 900

901 CGCCTGTAGGGATCATGGCCGGGAATGGCCTCAGACGCTTTGAGTGCCACCTTGAGGTT 960  
M A S D A L S A H L E V  
(orf32)

961 TCCGACTGTGCGGACGCGCGGGGGGATCACGGTGACGAATGACGGATCTGAACTCGCCGGG 1020  
S D C R Q R G G I T V T N D G S E L A G

1021 CAAAACGTGGCGGCGGTCCGCTTCGAGCGGTATTCCGCGATCGCGCCGAGCGGACCGCC 1080  
Q N V A A V R F E R Y S A I A P E R T A

1081 ATCTGCACAAAGGTGCCGCGACCGGTTACGACGAGCTCAACCGCGGGCCGAGCTGACA 1140  
I L H K G A A T G Y D E L N R R A E L T

1141 GCCACGCGCTGGCGGACGCGGGCGCCGCCCTCGACCCTGGTGGCAGTGGCCCTCCCA 1200  
A T R L A D A G A G P S T L V A V A L P

1201 CGCGATCCCGACCTCGTCGCCACCCTGTGCGCCCTGCTCAA CTGGGTGCCGATGCCTT 1260  
R D P D L V A T L C A L L K L G A A C L

1261 CCCCTGGATCCCGGCATACCGGCCGGGCGGTGCGCGAGATCATGGCCGACGCGTCCCCC 1320  
P L D P G I P A G R L R E I M A D A S P

1321 GACGTTCTCGTCACCACCCGTGCCGTGCTCCGGCATTACCGGTGACGGACCGTCCTT 1380  
D V L V T T R A V A P A F T G D G P V L

1381 TTCCTGGACGACGCTCCTCCGACCTGCTCCGCCGTCCTTCCACGGCACTCAGCGGGGACC 1440  
F L D D A P P T C S A V L P R H S A G T

1441 GCGTCGGAATCGCCTATGTGCTGTACCCGACGACTCCTGACGAGAAGTCCGAAAATTCG 1500  
A S E I A Y V L Y P T T P D E K S E N S

1501 GTCGTCTCCTATCGTGATATGGCGCGCTACCTTGACGACCCCACTGCCGGGATTCCGGCG 1560  
V V S Y R D M A R Y L D D P T A G I P A

1561 AGGGCGGAGATTCTCCGGCTGGTCGCGCCGCTCCTGTCCGGCGGTCTGTGGTGCTGGAC 1620  
R A E I L R L V A P L L S G G R L V L D

1621 GCCGACGAGACCCGCGCCCGCGGTACCCGTGAGGCGCCGCGACATGGTGGAGGAC 1680  
A D E T R P R P V T R E A P R D M V E D

1681 GTCGTGGCGCAGGTCTGGTGCGCCGTCTCGGCGTGGACCGGTGGGCGTGGGGACCGC 1740  
V V A Q V W C A V L G V D R V G V R D R

1741 TTCTTCGACCTGGGCGGCAAGTCGCTGGCGCGGTCCAGGTGGTGGCGGCCTGCGGAAG 1800  
F F D L G G K S L A A V Q V V A R L R K

1801 CTGCTCGGCGTCGAGCTGCCGCTGCGGGCCCTGTTTCGACGCGCCGACGGTCGAGGAGCTG 1860  
L L G V E L P L R A L F D A P T V E E L

1861 GCCGCCCGGTGCGGGCCGAACAGGCCGGCGCCAGGGCGTCCGGGAGGAGGCGCGCTC 1920  
A A R V R A E Q A G G Q G V R E E A A L

1921 GAGCCGGTGGGCGGAGCGAGCCGCTGCCGCTGTCGTCGACAGCAACGCCTGTGGTTC 1980  
E P V G R S E P L P L S F A Q Q R L W F

1981 CTGGACCGCTTGATGCCCCACCGCGCCTTCTACACGATGTGCGACGCGTCCGCGTCCGG 2040  
L D R L M P D R A F Y T M C D A F R V R

2041 GGCGGGATCGACCTGGGTGCGCTGCGGGCGGCCCTGCGGATGCTGGTGGGACGGCACGAG 2100  
G G I D L G A L R R A L R M L V G R H E

2101 ACGCTGCGGACGGCGTTCGTCGAGCGGGACGGTGTGCCGTACCAGCTCGTCCGGTCCGGCC 2160  
T L R T A F V E R D G V P Y Q L V G P A

2161 GACGGGCCCCGTGCGGGCGCGTGGCCGCTCCACGCGGGTCGACCTGTGCTGTGGAG 2220  
D G P G A R R V A A P T R V D L S L L E

2221 CCCGCCGAGCGGGAGGAGGCGGTGCGGAACCTGGTGGCGGGGAGGCGGGACCCCGTTC 2280  
P A E R E E A V R N L V A A E A R T P F

2281 CGGCCGGCGGACGGCGCGCTGTGCGCGTGGTGGTGGCCCGGTGGCGGACGATGATCAC 2340  
R P A D G A L L R V V V A R L A D D D H

2341 GTGCTGGTGGTCAGCACGCACCACATCGTCTCCGACGCCTGGTCCGTGGGTGTGCTGGTG 2400  
V L V V S T H H I V S D A W S V G V L V

2401 GACGAACTCGGACGGCTGTACCGCGAGTGCCTACCGGAGATCCCGCCGCTGCCCCCG 2460  
D E L G R L Y R E C V T G D P A A L P P

2461 CCGGCCGTCCAGTACGCCGACTTCGCGGTCTGGCAGCGGGCCTGGATGGCCGGTCCGGTG 2520  
P A V Q Y A D F A V W Q R A W M A G P V

2521 CAGGAGGAGCATCTCGGTACTGGAAGCGGGCCTTGGACGGCGCTCCCTCGGTGCTGCGG 2580  
Q E E H L A Y W K R A L D G A P S V L R

2581 CTGCCCATGGACCACCGCGGCCCGGTGTCAGTCCGAGCGGGGCGAGACGGTCCGGTTC 2640

43

3841 CAGGCCACCTTTCTCGGCCACGACCCGTATCTCGCCGGGGCCGACGGCGTACCGCCCGGG 3900  
Q A T F L G H D P Y L A G A D G V P P G

3901 GACGCGAAGTGCCTACGACGCTCACCGCGCCCTTCACGTTTCGACGCGTCCATGGAGCAA 3960  
D A K L R T T L T A P F T F D A S M E Q

3961 CTGAGCTGGATGCTGGCCGGTACGAGCTGTTTCATCGTGCCCGAGGACGTGCGGCGCGAC 4020  
L S W M L A G H E L F I V P E D V R R D

4021 CCCTCGGCGCTGGTCCGGTTCGTCCGGGAGCACCGGATCGACGTCATCGACACGACCTCC 4080  
P S A L V R F V R E H R I D V I D T T S

4081 TCGCAGCTCGAACTCCTCGTATCGCACGGGCTGTTGGACGGAGAGTGGGCGCCGTCATG 4140  
S Q L E L L V S H G L L D G E W A P S M

4141 GTCATGGTGGGTGGCGAGGCGGTCTCGCCGTCGCTGTGGCGGACCTTGCGGGACCAGCGG 4200  
V M V G G E A V S P S L W R T L R D Q R

4201 CGCACTCGCTGTTTCAACCTGTACGGGCCTACGGAGGCGACGGTCGACGCCACCTGCCAC 4260  
R T R C F N L Y G P T E A T V D A T C H

4261 GACCTGTCCGACCCCGCGACGTCGCCGTCATCGGCACCCCACTCCCCACACCCACGTC 4320  
D L S D P A D V P V I G T P L P H T H V

4321 CGCGTGCTCGACGACCGACTGCGACCCGTACCCGTGGGCGTCGCCGGCGAGATCTACCTC 4380  
R V L D D R L R P V P V G V A G E I Y L

4381 GGCGGAACCGGCTGGCCCGCGGCTACCTCAACCGCCCGCCCTACCGCCCGACGCTTC 4440  
G G T G L A R G Y L N R P A L T A R R F

4441 GTCGCCGACCCCTACCCGACACCCCGGCGAGCCGCTGTACCGCACCGGCGACCGCGCC 4500  
V A D P Y P D T P G S R L Y R T G D R A

4501 CGCTGGCGCCCCGACGGCACCCCTCGAATACCTGGGACGCACCGACGACCAATCAAGATC 4560  
R W R P D G T L E Y L G R T D D Q I K I

4561 CGCGGCTTCGCGCTCGAACCCGGCGAAATCGAGGCCGTCTCACCACCAACCCCGCCGTC 4620  
R G F R V E P G E I E A V L T H H P A V

4621 AAGGAAGCCGCGCTCGTCGACGACGCGCACGCGGCTGGTTCGCTACGTCACGCTCGCG 4680  
K E A A V V D D A H A R L V A Y V T L A

4681 GAAGCGGGCGCGCCGGCCCCACCGACGTACGCCGGTTCGCGCAGGGGCGGCTGCCCGCC 4740  
E G G G A G P T D V R R F A Q G R L P A

4741 CACATGGTGCCGTCGCGGGTGGTCTCTGGAGGCGCTGCCACTGACGTCGAACGGAAG 4800  
H M V P S A V V V L E A L P L T S N G K

4801 CTGGACCGCGCGCCTGCCGGCGCCCGCGGGCAGACCGGAAGTGGATGTCCGCTTC 4860  
L D R A R L P A P A A G R P E L D V R F

4861 GTGGCGCCGCGGACATGGTGGAGGAGTTCGTGGCGCAGGTCTGGTGCGCCGTGCTGGGC 4920  
V A P R D M V E E V V A Q V W C A V L G

4921 GTCGACCGGTTCGGTGTGCACGACGACTTCTTCGAGCTGGGCGGGCACTCGTTGCTGGTG 4980  
V D R V G V H D D F F E L G G H S L L V

4981 GTCCAGGTGATGACCCGGATACGAAAGCTGCTCGGCGTCGAGGTGCCGTTGCGGGAGCTG 5040  
V Q V M T R I R K L L G V E V P L R E L

5041 TTCGACGCCGCGACGGTCGAGGAGCTCGCCGCCCGCTCCGCGCCGACGGACCGAGGGC 5100

**Figure 6.** The effect of the initial concentration of the monomer ( $C_0$ ) on the polymerization rate at different temperatures. The reaction conditions were as follows:  $[AIBN] = 0.005 \text{ mol/L}$ ,  $[M] = 0.05 \text{ mol/L}$ ,  $[KBrO_3] = 0.005 \text{ mol/L}$ ,  $[H_2SO_4] = 0.005 \text{ mol/L}$ ,  $[NaNO_2] = 0.005 \text{ mol/L}$ ,  $[K_2S_2O_8] = 0.005 \text{ mol/L}$ ,  $[K_2Cr_2O_7] = 0.005 \text{ mol/L}$ ,  $[K_2FeO_4] = 0.005 \text{ mol/L}$ ,  $[K_2CoCl_4] = 0.005 \text{ mol/L}$ ,  $[K_2NiCl_4] = 0.005 \text{ mol/L}$ ,  $[K_2CuCl_4] = 0.005 \text{ mol/L}$ ,  $[K_2HgCl_4] = 0.005 \text{ mol/L}$ ,  $[K_2PtCl_4] = 0.005 \text{ mol/L}$ ,  $[K_2AuCl_4] = 0.005 \text{ mol/L}$ .

45

6301	TCCCTCGACGTCGTCTCGGGCGGGCGGGGAAGCGGTGCGTGCTGACGTACTGTCCCAGC	6360
	S L D V V S G R R G K R C V L T Y C P D	
6361	CTGTTTCGACCGGCCCCGCATGGAGGTGCTGGCCGGCCACTACCTGACCTGCTCGGCGCG	6420
	L F D R P R M E V L A G H Y L T L L G A	
6421	GCGGCCGACGATCCCGGTCTCCGCGTCGGCGACCTCCCGCTGAGCGACGACGTCGAACGC	6480
	A A D D P G L R V G D L P L S D D V E R	
6481	CTGCGCCTGCTGGGCGGGTCCCGCCCGCGGTACCTGCCCGCGCCCGGGCGGAGACCGTC	6540
	L R L L G G S R P R Y L P A P G A E T V	
6541	CCTGACGCCTTCGCGCGCAGGTGCGGGCGACACCGGACGCGCCCGCGTGGTCCACGGG	6600
	P D A F A A Q V R A T P D A P A L V H G	
6601	GACTCGACGCTGACGTTCGCCGAGCTGGACACCCGGGTACCGCCCTGGCCGTGCGGTTG	6660
	D S T L T F A E L D T R V T A L A V R L	
6661	CGGCGCTGCGGCGTGGCCGCCGAGACGCCGGTTCGCGGTGTGCCTGCCGCGCTCCGCCGAC	6720
	R R C G V A A E T P V A V C L P R S A D	
6721	GCCGTCGTGGCCCTCTCGGCGTCTCGGGCGGGCGGCGTCTATGTGCCAGTGGATCCG	6780
	A V V A L L A V L R A G G V Y V P V D P	
6781	GAGTGGCCCTCCGGCCGCGTCGCCACGTCTCGACGAGACCGCGGCCCCCGTCGTCATC	6840
	E W P S G R V A H V L D E T A A P V V I	
6841	ACCCGCGACCTGCCCGCCGATCCCGCCGCGTCCACCTCGACCCGCGCCAGGCCCGGCC	6900
	T R D L P A D P G R V H L D P R Q A P A	
6901	GACGACCGGATCCCTGCGCGCCTCCACCGCGACCAGGCCGCTACATCATCTTCACC	6960
	D D R D P L P R L H R D Q A A Y I I F T	
6961	TCGGGCTCCACCGCGCCCCCAAGGGCGTCTGTCGTCGACACGGCTCCCTGTACCACCTC	7020
	S G S T G A P K G V V V R H G S L Y H L	
7021	CTGGGCCACGTACGGCGCATGGCGGAGGGCGGCCCCCGGGAACGTGCGGCACACCACC	7080
	L G H V R R M A E G G P R R N V A H T T	
7081	GCGATGACCTTCGACCCGTCGCTGGAACAGTTCCTGTGGCTCGTCGCCGACACACCCTG	7140
	A M T F D P S L E Q F L W L V A G H T L	
7141	CACGTCGCGCCCGAGGAGGTGCGCCGCGATCCCGAGGCGCTGGTGGCCCTGGTGGGCGC	7200
	H V A P E E V R R D P E A L V A L V R R	
7201	GCCGCGATCGACGTCTCAACGTACCCCGTCCACCTGACCTGCTGATCGAGGCCGGG	7260
	A A I D V L N V T P S H L T L L I E A G	
7261	CTGCTGGAGGGCGACCGGGTCCCGGTACGGTCTGCTGGGTGGCGAGGCGGTGCCCGCG	7320
	L L E G D R V P G T V L V G G E A V P A	
7321	GCGTGTGGCGGACCTGCGCGAACGACGGGAGCCACCCGCTTCTTCAACCTGTACGGG	7380
	A L W R T L R E R T G A T R F F N L Y G	
7381	CCTACGGAGGCGACGCTGACGCCACCTGCCACGACCTGTCCGACCCCGCGACGTCCCC	7440
	P T E A T V D A T C H D L S D P A D V P	
7441	GTCATCGGCACCCCACTCCCCACACCCACGTCCGCGTGTCTGACGACCGACTGCGACCC	7500
	V I G T P L P H T H V R V L D D R L R P	
7501	GTACCCGTGGGCGTCGCCGGCGAAATCTACCTCGGCGGAACCGCCTGGCCCGCGGTAC	7560

47

8761	GTCGCCAACGACCGTCTGACGCGGGACAACGAGCTGCTCGACGCACCCGCCGCACGACG V A N D R L T R D N E L L D A P A R T T	8820
8821	GCCGTCGAGCCCGAGGACCTGTGGGGGCTGGCGGACTCCACCCCTACCGGGTGAGCGTC A V E P E D L W G L A D S T P Y R V S V	8880
8881	AGCTGGGCCCGCCCGATCCGCGGGCGCGATGGACGTCCTGCTGGTCCGGCGGGACGCC S W A A A D P R G A M D V L L V R R D A	8940
8941	CACGACGACGGTCCGCTGCTCGTCCCCACCCCGTACCGGAGCCCTCGGCACCGCTGACG H D D G P L L V P H P V P E P S A P L T	9000
9001	AACACGCCGACCCCGCACCCGTCCGCGCGGCAAGGGGGCTCGGCCGCGGACGGGTGCGT N T P T R H P S A R Q G G S A A D G L R	9060
9061	TCCTGGCTCGCCGAGCGGCTTCCCGCGCACCTGCTGCCCGCGAGGATCACCGAGGTGGAC S W L A E R L P A H L L P A R I T E V D	9120
9121	GCGCTGCCCCGCACCGGCACCGGCAAGCTCGACCGGGCGCGCTCGGCGGACTCGTGACC A L P R T G T G K L D R G A L G G L V T	9180
9181	GCGGGCCGTGGCGCCCGGGCGGCGACCGCCCCGCCACCGCCCCCGTACGGGTCTCGAA A G R G A R A G D R P A T A P R T G L E	9240
9241	CGGACCTGGCCGACGCGTGGGCGGGGTGCTCGGCCTCCCCGAAGTCGGCGTGACGAG R T L A D A W A R V L G L P E V G V H E	9300
9301	AACTTCTTCCCTCGGCGGCGACTCCCTCCTCGCCGTGAGGGCTGTGCCCCGGTGCCCG N F F A L G G D S L L A V R A V A R C R	9360
9361	CGTGCGGGGTCCGACTGACCGTCCGCGAGTTGCTGAGCGAGCAGACCGTCGCCGCGCTC R A G V R L T V R Q L L S E Q T V A A L	9420
9421	GCGGCGGCCCTCGAGGAGGAGTCTCAATGATGAAGTCAAGCCGCTTGCGCGACCGGCAGC A A A L E E E S Q * M M K S S R L R D R Q L (orf33)	9480
9481	TCGGGGGTGAAGACCCGGTTGTGCGCGAGGAGCCACAGGACGCTGGCCCGACGCCGT G G E D P V V A Q E S P Q D A G P T P C	9540
9541	GCCAGGGCGATGACGGCTTGAACGTGTTTGACGCCCTCGCCGCGCTTCTTGAGGTAGAAG Q G D D G L N V F A A L A A L L E V E V	9600
9601	TCCCGGTTCCGCCCCCTCCCGCATCATGCTGGTTTGGGCCGACATGTAGAACACTCGTCGC P V R P L P H H A G L G R H V E H S S Q	9660
9661	AGGCGGCGGCTGTAGCGCTTGGGCCGATGCAGGTTGCCAGTGCAGACCGGAGTCGCGG A A A V A L G P M Q V A S A T T G V A G	9720
9721	GGGACGGGCACCGAGCCGCGCCGAGGCCAGGTGACCGGCGTCGGCGTAGGCCGTGAGG D G H Q A G R R G Q V T G V G V G R E V	9780
9781	TCGCCGCGGCGACGACGAACTCGGGCCGAGGATCGGCCCATGCCCGGCAGAGACTCG A G G D D E L G A E D R P H A R Q R L D	9840
9841	ATGATCTCGGCCTGTGGATGGCTGCGGAACGTCTCGCGGATCTGCTGGTCAATCCGCTTC D L G L W M A A E R L A D L L V N P L Q	9900
9901	AGACGGTCGTCCAGGCCAGGATCTGCGCGGCCAGGTGAGCCACGATCTGGGCGGCGACG T V V Q G Q D L R G Q V S H D L G G D V	9960



9961 TCCTCCCCGGGCAGCGCGGTCTGCTGAGCCTGGGCAGCCTCCAGCGCCGTCGCGGCGACG 10020  
L P G Q R G L L S L G S L Q R R R G D G

10021 GCGTCGGCACCGCGCACGCCCTCGGTTGGCCAGCCAGGCCGTACGCCGGGCCCGGCCGCGG 10080  
V G T A H A S V G Q P G R Q P G P A A A

10081 CGGCGGAGAGCTCGCGGGGTCTGGTAGCCCGTCAGCAGGACCAGCGCGCCCTTCTGCGAG 10140  
A E S C R G L V A R Q Q D Q R A L L R A

10141 CTGTAGTCGAAGGCCCGTTCCAGCGCGGGGAAGACGCCGGTCAGCGTGTGCGGAGACGG 10200  
V V E G P F Q R G E D A G Q R V A E T V

10201 TTGATCATCTGACCCGGTCGGCCACGAGGTCGGAACGGTGGGCGGTCAGCAGCGCGAGG 10260  
D H P D P V G H E V G T V G G Q Q R E V

10261 TCGGCGGCCAGCTGGGCGGGCACGTGATCGACGCGAAGTCCCGTCGGTTGCGGGCGGTT 10320  
G G Q L G G H V D R R E V P S V A G G F

10321 TCGGCGATGACGTAGGCGTCCGCGGCGTCTTCGCCTCGCCCCGGTAAGCGCCGGAC 10380  
G D D V G V A G V G L R L A P V S A G H

10381 ATGCGGTTGACCGTGCGGCCGGGCACGTAGACGGCCTGCTGGCCGTGGGCCGCGAGCAGG 10440  
A V D R A A G H V D G L L A V G R E Q G

10441 GCCAGCAGCAGCGCGGAGGACGTGCCGAGATGTCCACTGCCAGTGGACCTCGTCGGCC 10500  
Q Q Q R G G R A G D V H C P V D L V G Q

10501 AGGTCGAGGATCTACCCATGGCGGTGAGGATCGCCGACTCATCGTTGCCGATCTTCTTC 10560  
V E D L T H G G Q D R R L I V A D L L R

10561 GACCACAGCCTCACACCGGTCTCGTCGACCACCGCCGCCAGTGATGCCCTTGCCCGCG 10620  
P Q R H T G L V D H R R P V M P L A R V

10621 TCGATCCCGGCCAGACCCGGGCGCGTCTCGCCACTCGCCCTCCTCACTCCGAACA 10680  
D P G P D P G P S L A H S P L L T P N S

10681 GCATCCCGTGCACCCGAGGAACACCCGCTGTCTATCTCCGTAAAAAGCGACCGAAGCGCA 10740  
I P S T R G T P R C H L R K K R P K R T

10741 CATCTCAATCAGCAGCCAGGGCGCCCCGAGAACCGGGCGGCCACTCCTGTAAAGCCACT 10800  
S Q S A A R A P R R T G R P L L V S H \*

10801 GACGGCAGAGAACCATAAGCCACACCGGCCCTCCCGGGCCGCTAACAACTTACGGAGA 10860

10861 ACCATGACTGACCTGCCGTTGCGTACCGTCGCACTACCGGTGAGGAGAGCGCGGAGGTC 10920  
M T D L P L R T V A L T G E E S A E V  
(orf34)

10921 GACGACCTGTGCGCACGCTGGCCGACGTGCCGGTCGACTCCACCGTGGGACTGCTGCAC 10980  
D D L L R T L A D V P V D S T V G L L H

10981 CGCACCCGGTTCGCCGACAGGAAGTCCGCTGCGCATCCGCGCGAGCTCACGGGGATG 11040  
R T R L A A Q E L P L R I R A E L T G M

11041 CGGCTCTACGACAGCCCGCGCGCCCTCGTCTGTCACGGGCTTCGGCGTCGACGACGAACGG 11100  
R L Y D S P R A L V V T G F G V D D E R

11101 ATCGGACCGACCCCCCGGCCCGTCCCGCCCCGGATCCCGAGCGGACCCGCGACCTGGAG 11160  
I G P T P A A R P A P D P E R T R D L E

11161 CTGCTGCTTTTGTGTCACGCGGCCCTGCTCGGCGAGGCGTTTCGGCTGGGCGACCCAGCAG 11220  
L L L L L H A A L L G E A F G W A T Q Q

11221 AACGGCCGGCTCGTCCACGACGTGCTGCCCGTTCCCGGTGAGGAGACCGCGCAGATGGGT 11280  
N G R L V H D V L P V P G E E T A Q M G

11281 TCCAGCAGCGAGACCGAGCTGCTGTGGCACACCGAGGACGCGTTCCACCGCTGCGCTGC 11340  
S S S E T E L L W H T E D A F H P L R C

11341 GACTACGTGGGCTGCTGTGCCTGCGCAACCACCAGCGCGCCGCGACCACCGTGGGCTGG 11400  
D Y V G L L C L R N H Q R A A T T V G W

11401 CCCGACCTGTCCCGGCTCACCACCGAGGACCGTGCCCGTCTCCTCGAACCCCGCTATCTG 11460  
P D L S R L T T E D R A V L L E P R Y L

11461 ATCCGCCCCGACACCTCGCACACGCCCGCGCAGAACGCGACGGGCACGCGGTCCGCCGAG 11520  
I R P D T S H T P A Q N A T G T R S A E

11521 CGTTTCGCGGCGATCGCCGAGATGGACGACGCCCCGGAGCGCGTCGCGGTCTCTGTTCGGC 11580  
R F A A I A E M D D A P E R V A V L F G

11581 GACCCCGAGGACCCGTACCTGCGGATCGACCCGGCCTACATGAGCCCGGCCCGGGGAC 11640  
D P E D P Y L R I D P A Y M S P A P G D

11641 GCGGCCGCCCGGGCGGTACGACACCGTCACCGCGCTCATCGAGGACGAGCTGCGGCAC 11700  
A A A R R A Y D T V T A L I E D E L R H

11701 GTCGTCCTGGACGCCGTTCACTGCTGCTGGTCGACAACTACCAGGCGGTGCACGGCCGC 11760  
V V L D A G S L L L V D N Y Q A V H G R

11761 AAGCCGTTCCGCCCGCCTACGACGGCCGCGACCGCTGGCTCAAACGCGTCAACATCACC 11820  
K P F A A A Y D G R D R W L K R V N I T

11821 CGCGACCTGCGCCGTTCCCGGTCCGCGCGGCGGTGCGGCACCTCGCTGCTGGTGTGAGGG 11880  
R D L R R S R S A R R S A T S L L V \*

11881 AGGCACCATGGATTTCCTCCCTCACCCGCGTCAACCCCTGGTTCAGCGGCGGCTGCGACGG 11940  
M D F P L T R V N P W F S G G C D G  
(orf35)

11941 CCGCCCCCGGTGCGGCTGTGCGCGCTGCCGTACGCGGGCGGCACCGCCGCGCTTCAA 12000  
R P R V R L C A L P Y A G G T A A V F K

12001 GGACTGGCCCCCGCGCTGCCCCCGGAGTGGAGCTGCTCACCGCACCTGCCGGGACG 12060  
D W P A A L P P G V E L L T A H L P G R

12061 CGGCGACCGGTTACCGAACCGCCCCCGGCCACCTGGAGGAGACCGCGAGCGGCTGTG 12120  
G D R F T E P P P A T L E E T A E R L C

12121 CGAGGCGCTGCCGCCGAGTGACCTGCCCACGGTCGTCCTCGGCCACAGCATGGGCGCCCT 12180  
E A L P P S D L P T V V L G H S M G A L

12181 GCTGGGGTACGAAGTGGCGGCGGGCTCGCGGCGGGGCCGCGCCCCAACCTGCTGAT 12240  
L G Y E V A A R L A A R G R A P N L L I

12241 CGCCGCGGCTGCGCTCCCCCGCACGTTCCGCGGACGCTCCGGTCCGGTGACCGAGGC 12300  
A A A C R P P H V P P D A S G P V T E A

12301 CGAGCTGGCGGCCACCTGCGGGCCGAACGCCCATGGGACACGGCCCTGAGGGACGAGGA 12360  
E L A A T L R A E R P W D T A L R D E E

12361 ACTGATGGAAGCGGTGCTGCCCCCCTGGTCGCCGACATCACGGCCGGCGACCGCTACCA 12420  
L M E A V L P A L V A D I T A G D R Y H

12421 CCGCCCGCGGCCCCCGCGCTCGACCTCCCGTGAAGGTCTACATCGGCGCCGACGACGA 12480

R P R P R P L D L P L K V Y I G A D D D

12481	CGGCACCGACTGGCGCACCACCTGGGCTGGCGCGCGTGCACCGCCCGGGACTGCGAGGT G T D W R T T L G W R A C T A R D C E V	12540
12541	CGTCGTCCTGCCCCGGCGGCCACTACTTCCTGGAGACCGACCGCGCGGCCGTCTCACCCG V V L P G G H Y F L E T D R A A V L T R	12600
12601	CGTCGCCACGGACCTCGCCGAAGCCGAGGTAGGGGCATGACCGCGCGCTCGACGCCACA V A T D L A E A E V G A * M T A R V D A T (orf36)	12660
12661	CCCACCTACCTGGCGGTGCTGGCGGTGCGCGAGGCCCGCGCCCCGCTCCTCGGCAGCTGC P T Y L A V L A V R E A R A P L L G S C	12720
12721	CTGCCCCGATGTCTTCGCGGTGCTGCCGCTCGCCCTGCTGTGTCTCGGTCCGGGACGCG L A R M S F A V L P L A L L L S V R D A	12780
12781	ACGGGGTCGTTTCGCGCTCGCCGACTGACCTCCGGCGCGCTGTCTCGGCCACGCTCACGCTG T G S F A V A G L T S G A L S A T L T L	12840
12841	TTCGCGCCCGCCCGCGCCCGCTGATCGACCGCCGGGGCTCACGGTCCGGACTGGTCCGG F A P A R A R L I D R R G S R S G L V R	12900
12901	CTGACCGTCCCGTACCTGCTGGGGCTCGCCGTGCTGATCACATTGGCCGAGGCGGAAGCG L T V P Y L L G L A V L I T L A E A E A	12960
12961	CCCACCGCGCGCTGCTCGTCGCGCCCGCGGTGCGGGCGTGTTCGCGCCCGCGCTCGGT P T A A L L V A A A V A G V F A P P L G	13020
13021	CCGACCATGCGCGTGTGTGGGCGAGGATCCTGCACGGCCGTCAGCCCTGCTGCACACC P T M R V L W A R I L H G R Q P L L H T	13080
13081	GCCTACGCCCTCGACTCCGTACCGAGGAGTGGTCTTCACCGTGGGGCGCTGCTGGCG A Y A L D S V T E E V V F T V G P L L A	13140
13141	GGCGGCCTGATCGCGGTGCGGGCACCGCTCGCGTCGATGATCACGGTCATGGTGTGATC G G L I A V A A P L A S M I T V M V L I	13200
13201	GCGGCCGCTACCGCTGCTTCGTGCTGTCCGCCGCGACCGCCCGCCCCCGCTCGGGC A A G T A C F V L S A A T A A A P A S G	13260
13261	GAAGCCGACGAGGACCGGCCGACGCGCGGCCCATGGTCTGCCCCGGATGCGCACGATC E A D E D R P H G R P M A L P G M R T I	13320
13321	GTGCTGTCTTCGCGCGCGTGGCCCTGGTGTGCGGGGTGCTCCAGGTCGTCTGCCGTTT V L S F G G V G L V V G V L Q V V L P F	13380
13381	ATCGCCGACACGCGGGCTCGCCCGCGCGGCGGCATCCTGCTGTCCATGCTGTGCGCG I A D H A G S P G A G G I L L S M L S A	13440
13441	GGCAGCGCGGTGCGCGGCCCTCGCCTACGGGCGGATCGCCTGGCGCTCGACGCCCGTGGC G S A V G G L A Y G R I A W R S T P V R	13500
13501	CGGTTCGTGGTGTCTGTCACCGGGTTACAGCTGGCGGTGCTGCGCTGTGCTGACCGCG R F V V L V T G F T L A V L P L C L T A	13560
13561	AGCCCGGTGCGGGCGGGGCTTCGCCCTCCTCGTGGGACTCTGCCTCGCCCCGCTGTTT S P V P A G A F A L L V G L C L A P L F	13620
13621	ACCACCGCCTACCTGCTGGTCAACGACCTGGTGACGGCGTGGGGACCGCACCCACCGAG T T A Y L L V N D L V T A S G T A P T E	13680

13681 GCCAACACCTGGGTCTCCACGGCCAATAACGGAGGGTTGCGCGGGGAGCGCCGCGCC 13740  
A N T W V S T A N N G G F A A G S A A A

13741 GGTGTGCTGCTCGACTCCCGGGGCCCCACCGTCACCGTCACCGCGCGTTCGCGGTGCGC 13800  
G V L L D S R G P T V T V T A A F A V A

13801 GCCGCGACCGCCGTATGACCGTTCTGCGCCGCGGACCCTGCTCCTCGGCGCCGGACAC 13860  
A A T A V M T V L R R R T L L L G A G H

13861 CCCGAACCGGCGCGGCCACACCGCGGACCGCACCGCGCGAAGCCGAGGAGTGA 13920  
P E P A A A T P A D R T A P A E A E E \*

13921 ACCGATCGTGTCCAAGAACGCGCGCACTGGTCGCGCATCCGCACAGGGGACGCCCCCGG 13980  
M S K N A A H W S R I R T G D A P G  
(orf37)

13981 CGTCGTA CTGCGCGTGGACTTCTACGGAACGGGCGCCAGGAAGCCACCTTCCGCCACCT 14040  
V V L A V D F Y G T G R Q E A T F R H L

14041 GTGCGACCTGCTCACGGATCCGGTCGAGGTCTGGCACGCGGTCCCGCCCCCGCCGACGG 14100  
C D L L T D P V E V W H A V P P A P D G

14101 CGACTGGTCCACGGCCACCGGCGCGGTCACTGCGCTGGTGACCGAGGGGCTCGACAC 14160  
D W S T A T G A G H L R W W T E G L D T

14161 GGTCTCGCGGGACGGCCGGTGC GGCCCTCGTCGGCTACTGCGCGGGCGGCGTCTTCGC 14220  
V L A G R P V R A L V G Y C A G G V F A

14221 CTCGGCCCTCGCCGACGCCCTCGTCGAACGGGAGGGCCACCGCGCGGGTCTGTGTGT 14280  
S A L A D A L V E R E G H R P R V V L F

14281 CAACCCAGCGCGCCCGGCGTGCACGCTCACCGCGACTTCCGCGGTCTGATCGCCGG 14340  
N P S A P G V A T L T R D F R G L I A G

14341 CATGGACCTCCTCACGGACGGGAACGCGCCGCTCTGCTGGCCGAGACGACCGGATCCG 14400  
M D L L T D G E R A A L L A E T T A I R

14401 GCGGGCACACGCCCCCGACGCGTGGTACCGGTGCGCGAAGCTACGCGCCCTGTACCG 14460  
R A H A P D A L V P V A E R Y A A L Y R

14461 CGAGGGCTGCGACCTCCTGTGCGAGCGGCTCGGCGTGGACGCCTCCTTCGGCGCCGAAC 14520  
E G C D L L C E R L G V D A S F G A E L

14521 GGCCGCGGTCTCCACTCCTACCTGGCCTACCTCACGGCGGCGCTCGACGTGCCCCCAC 14580  
A A V L H S Y L A Y L T A A L D V P P T

14581 CCCGCTGTGGCGGGCGCGTCTCGCTCACCTCCCGGAGCACCAGGGCACCGACTTCAC 14640  
P L W R G A V S L T S R E H Q G T D F T

14641 CGACGTCGAGCACGGCTTCGACGTCGCGGTGCGGAACGCTGAGCTCCCCCAGGTCGT 14700  
D V E H G F D V A R A E L L S S P Q V V

14701 CGCGGCGCTGACCGCGTCTCCTCGCGAACACGAGGCGAGCCGATGACCCTCACCTGCGG 14760  
A A L T A L L R E H E A S R \*  
M T L T L R  
(orf38)

14761 GACGCCTTCTCGACAGGCGCGCGGACCGCCGACGCGCGTCTGACACGCGAC 14820  
D A F L D Q A A R T P D A H A V V H G D

14821 ACTGTATGGACGTACCGGAACGGAACGCGGGCGGCGCATGGCCCGGACGCTGGCC 14880  
T V W T Y R E L E L R A G R M A R T L A

14881	GCACGCGGCGCGGGCCCCGGGCACGCTGGTGGCGGTACGCCTGCCGCGCGGTCCCCGAACCG A R G A G P G T L V A V R L P R G P E P	14940
14941	GTCGCGCGCTCCTCGCGGTCTGTGCTGACGGGAGCGGGCTACGTGCCGCTCGCCGACGAC V A A L L A V V L T G A G Y V P L A D D	15000
15001	GACCCGCCGACCGGTGCCGGCACATCCTCGACGACTGCGCCGCGCGCTGCTGTGGCC D P P D R C R H I L D D C A A A L L L A	15060
15061	GAGCACCCCTCGCGGGACGGACGCACCTCACCCCGGACGAGGCGCTGGCACCCGCCCGC E H P S R D G R T L T P D E A L A P A R	15120
15121	CCGTTCGACGCGGCCCCGGTGCGGGCCGGGACCCGGCGTACGTGATCTACACCTCCGGC P F D A A P V R A G D P A Y V I Y T S G	15180
15181	TCCAGTGGCCGTCCGAAGGGCGTGTGGTGAACAGGGCGCGCTCGGCGCCTACCTGGCA S S G R P K G V L V E Q G A L G A Y L A	15240
15241	CAGGCCCGCGCGCGCTACGACGGGCTGTCCGGACGGACGGTGTGCACTCCTCGCTGTCC Q A R A R Y D G L S G R T V L H S S L S	15300
15301	TTGACATGGCCGTGACCACTGTGTGGGGCCCGCTCGTCAGCGCGGCGCGATCCACGTG F D M A V T S L W G P L V S G G A I H V	15360
15361	CTCGACCTGAAGGCGATCGCCTCCGGCACCCAGCCCGCGCCCGCGCCTCGGCACGTCCG L D L K A I A S G T Q P P P A A S A R P	15420
15421	TCCTTCCTCAAGGTCACTCCGTCCCACCTGCCGCTGTGGGCCTGCTGCCGGACTCCTGC S F L K V T P S H L P L L G L L P D S C	15480
15481	CTGCCACCGGCAACTCGTGATCGGGCGGAGGCGCTGACCGGTCCGCGCTCGGACCC L P T G Q L V I G G E A L T G S A L G P	15540
15541	TGGCGCGCGCGCACCCGACGTACGGTCTGTAACGAGTACGGGCCCACCGAGGCGACC W R A A H P D V T V V N E Y G P T E A T	15600
15601	GTCGGCTGCTGCGCGTACACCGTCCGCCCGCGGTGACGCGGTGGACCCGGGTGCCGTCCCC V G C C A Y T V R P G D A V D P G A V P	15660
15661	ATCGGACGGCCGTTTCGCGGGCACCCGCCTGTACGTGCTCGACGCGGACGGCGAGCCGGTC I G R P F A G T R L Y V L D A D G E P V	15720
15721	GCCGTGGGCGGTGTGGGTGAAGTGCACATCGCGGGCGACCACTGGCGCGCGGATACCTG A V G G V G E L H I A G D Q L A R G Y L	15780
15781	GGGCGCCCGCGGTGACCGAGGAACGCTTCGTCCCGGACCCGTTCCGCGCGACGGCTCC G R P R L T E E R F V P D P F A A D G S	15840
15841	CGGATGTACCGCACCGGCGACCTGGTGCGCGAACGCCCGACGGCGACCTGGAGTACCTC R M Y R T G D L V R E R P D G D L E Y L	15900
15901	GGGCGCGCGGACGGGCGAGGTGAAGGTCTCCGGGTACCGGATCGAGCCCGCGAGATCGAG G R A D G Q V K V S G Y R I E P G E I E	15960
15961	GCCGTGCTCCGCGGCCACGCGGGGTGAGGGACTGCGCGGTGTCGCGCTCGGCGAGGCG A V L R G H A G V R D C A V V A V G E A	16020
16021	GACGCCCGCGGCTCGTCGCCTACGTGGTACCGGACCCGGACTCCCCGCCCGGCACCGCC D A R R L V A Y V V P D P D S P P G T A	16080
16081	GCGCCGCGCGGCACGCGGCCGAGGCGCTGCCCGCTACATGGTGCCGGCGACGTTGCTC 16140	

54

17341 TCGCGCTGCCGTGGCGAGAGGGCGACATCATGCTGGTCGACAACCTGAGGATGGCCCACG 17400  
A L P W R E G D I M L V D N L R M A H G

17401 GCCGCGAGCCCTTCACCGGCGAGCGCGCTACTCGTCGCGATGACCTCGGCGGACTCAT 17460  
R E P F T G E R R V L V A M T S A D S \*

17461 GAGCCGTGCCGACGCATCGGCACGCCGTCTCCCGTCGGGGCGCTACCATCGCCGCTGTC 17520

17521 TCGGCCATCACCACCCCGGGCGGAGGCAACCGGCCGTGCACATCCCGCCGTGGTCGCC 17580

17581 ACGGCACGCGCGATCACCCGCGCCATGACCGCCAGCCCGTTGTACATCTGCGGAGGCG 17640

17641 CCGCGATGACAGAGGTCCGAGGTGAAGTATCCGGGCGCTGCCGGGTGTGCTGGAGGCGC 17700  
M T E V R G E L I R A L P G V L E A R  
(orf40)

17701 GTGCGGCGCGGGCGGGGCACACGACCGCCTTCCTCGACGCACGACGGTGTGTACGTACC 17760  
A A R A G H T T A F L D A R R C V T Y R

17761 GGGAGTTGGAGCGCGCACCCCGCGGTGGCGGGGTACCTGGTGCGGTGGGGGTGCGC 17820  
E L E A R T R R L A G S P G A V G G A Q

17821 AGGGGAGACCGGGTGGCGCTCGTCAATGGGCAACCGGGTGGAGATGGCGGAGGGTTCC 17880  
G Q T G W R S S M G N R G G D G G G F P

17881 CTCCCGGTGCTGCGGGCGGAGCGGTAGGGGTGCCGCTCGATTCCGGGGCCACGGACGC 17940  
P R C C G P E R \* G C R S I P G P R T R

17941 GGAGCTCGCGTACTTCTCGACGACTGTGGAGCGGTGGCGGTGGTCACCGAGGAGACGCT 18000  
S S R T S S T T V E R W R W S P R R R C

18001 GCTGCCGCGGTCTCGCGATCGCGGGCGTACGGATCCTGGTGGGGGGTTCGGACGCCGT 18060  
C R G S R D R R A Y G S W W G V R T P S

18061 CCCGAGGGAGCGGTGCCGGCATCCACTCCTTCGAGCGGCTCGCGGCGTTCGGATCCGGG 18120  
R R E R L P A S T P S S G S R R R I R G

18121 GTGCGCGCCACGGGACGACCTCGGCCTCGACGAGCCGGCCTGGATCCTCTACACGTCGGG 18180  
A R H G T T S A S T S R P G S S T R R G

18181 GACCACGGGGCGGAGCAAGGGCGTGGTCTGCGGCCAGCGCGCCGCGTGTGGTCCGTGGC 18240  
P R A G A R A W S A A S A P R C G P W R

18241 GCGCGGTACGTGCCGTCTGGGGTCTGGGGCCGAGGACCGGTGTTGTGGCCGCTGCC 18300  
R R T C R R G V W G R R T G C C G R C P

18301 CATGTTCCACGCCTACGCGCACTCGCTGTGCCTGCTCGGGGTGGTGGCCGTGGGCGCGAG 18360  
C S T P T R T R C A C S G W W P W A R A

18361 CGCGTACCTCCTCGACCGGGCGCGAGCGTCTCGGGCGCTTGAGGAACAGCGGTGCAG 18420  
R T S S T G A R A S S G R L R N S G A A

18421 CGTCGTGGCCGGTGTACCCGCCACCTACCGCCTGCTCACGAGCGCCTTCCGCGACGCCC 18480  
S W P V Y P P P T A C S R A P S A T P P

18481 CCGGCCACCGCGCGGCTGCGACTGTGCGTCACCGGGGCTGCGCCGTGCCCGCGGGG 18540  
G H R P A C D C A S P G A A P C P P G L

18541 TGCGGGCGGACGTTGAGGAGCTGCTGGGCGTCCCGCTGCTCGACGTTACGGCAGTACCG 18600  
R A D V E E L L G V P L L D G Y G S T E

18601 AGACCTGCGGCAAGATCACGGTTGAGCGGCTCGGCGGCTCCCGGAGGGCGGTTGCCGGG 18660  
T C G K I T V E R L G G S R E G G C R

SEQ ID NO: 3 BLM gene PPTase ORFS 41

1 GGATCCTGCCCTACCCGGAATTCGCCAGTGGTGCGGCACCGAGCTCACGCCGACTGGCACGTCCGCTTCCGGGCGGCC 80  
81 GCCGCGGTCTACGGGCATCTGCACATCCCCCGGTGACCCGGTACGACGGCGTCCGCTTCGAGGAGGTGTGGTCCGGCTA 160  
161 CCCGCGGAGTGGCGGCCCGGCCGCCCGCGAGCCGCTCCGGCAGATCCTGCCCCAGCCCGTCGACGAGCCGGGAGCCC 240  
241 TCTGGTGATCGCCGCCCTCCTGCCCTCCTGGGCCGTCAACGAACACGCCTTCACCGACGCCCCGGACGACCCGGTGAGCC 320  
1 M I A A L L P S W A V T E H A F T D A P D D P V S L 26  
321 TCCTCTTCCCGAGGAGGCCGCCACGTCGCCCGCGCGTCCCAAGCGCTGCACGAGTTCGCCACCGTCCGGGTGTGC 400  
27 L F P E E A A H V A R A V P K R L H E F A T V R V C 52  
401 GCCCGCGCCGCCCTCGGCCGCTGGGCCTCCCGCCCGTCCGCTGCTGCCCGGCCGACGGGGCGCCGAGCTGGCCGGA 480  
53 A R A A L G R L G L P P G P L L P G R R G A P S W P D 79  
481 CGGGGTGGTGGGAGCATGACGCACTGTCAAGGCTTCGGGGCGCCGCGTCCGCCGGGCGCCGACGCCCGTCCGCTCG 560  
80 G V V G S M T H C Q G F R G A A V A R A A D A A S L G 106  
561 GGATAGACGCCGAGCCGAACGGGCCCTCCCGACGGCGTCTCGCCATGGTCTCGCTGCGTCCGAGCGCGAGTGGCTC 640  
107 I D A E P N G P L P D G V L A M V S L P S E R E W L 132  
641 GCCGGAATGGCGGCCCGCCGCGCGGACGTGCACTGGGACCGGTGCTGTTAGCGCCAAGGAGAGCGTCTCAAGGCGTG 720  
133 A G L A A R R P D V H W D R L L F S A K E S V F K A W 159  
721 GTACCCGCTGACCGCCCTGGAGCTGGACTTCGACGAGCGCGAGCTGGCGGTTCGATCCGGACGCCGGGACGTTACAGGCC 800  
160 Y P L T G L E L D F D E A E L A V D P D A G T F T A R 186  
801 GGCTGCTGGTGCCGGGACCGTGGTTCGGCGCGCGTCCGCTGGACGGGTTCGAGGGGCGTGGGCGCGGGGAGGGCCTC 880  
187 L L V P G P V V G G R R L D G F E G R W A A G E G L 212  
881 GTCGTACGCCCATCGCCGTGCGCGGCCCGCGCGTACCGCGAGGAATCGGCGGAAGGGCCGGAAGGAGCGACTGC 960  
213 V V T A I A V A A P A G T A E E S A E G A G K E A T A 239  
961 GGACGACCGGACCGCGTCCCGTAACCGCCCCGAACACCGCGTGGCGCCCGCGACCGTGTGCGGGGCGCCACGAACG 1040  
240 D D R T A V P \* 247  
1041 GGCGCCGGCCCGCGCGGCCCTCCGCCGTGCGGAGCGGAGGCCCGCGCGGACGCGCCCGTGTCTGTCGGATACGTGCGTC 1120  
1121 AGTCGGCGACGACAGCGTTGCGGTTGGTTCGAGTTGAGCAGCCGACGATGTCGATGGTGTGCGCGAGAGTTGATGGGG 1200  
1201 ATGTGGACGGGGATCTGGATGACGTTGCCCGAGACGACGCCCGGGGAGCCGACGGCCGCCCTTGGCGTTCGAGTCGGC 1280  
1281 GAGGGCGGTGCCGAGACGCCGCGGAGCGCCGTGCCCAAGGTGGCGGTGAGGGCCGCTGCCTTGGCGATTCTGTGACATGG 1360  
1361 GGTGACACGTTCTGTTCCGTTGACAGGGTTCGAGCTCACGGCTCTGACGGCCGGGAGCCCGGATCAACGCCCGATCACCC 1440  
1441 CGAAGGTTTGAATCGTGCGCGGACGGGTGACCGCGGCCGAACGGCCTCGCCGGGCCCCCGGAAGGTGCCATGACGTC 1520  
1521 CGTGCGCATCTGTACAGCCCGTCCCGCGCGCGTACAAGGACGACGACGCGCGGTGGACGGACGACCGCGCGGGGA 1600  
1601 GGGGAGGCCATGAGCCGATCGCGATCGTGGGGCGGGTCAGGCCGACTGCATCTGGCGCTGGGGCTGCTGGGGGCGGG 1680  
1681 GAGCGGCTCTTCCCGTCACGAGGTGCTGCTCGTGTCCGACGGGACGCCGACGAGATCCGCGCGGGCGGGTGCGGTGCA 1760  
1761 C 1761